OM EDITORIAL

AUTORES
Alarcón Anco Ronald Jesús
Gallegos Ruiz Conejo Ada Lucía

INTELIGENCIA ARTIFICIAL APLICADA A LA INVESTIGACIÓN CIENTÍFICA

Herramientas, ética y nuevas formas de producir conocimiento





Inteligencia artificial aplicada a la investigación científica: herramientas, ética y nuevas formas de producir conocimiento.



Autores:

Alarcón Anco, Ronald Jesús

Universidad Nacional Mayor de San Marcos

Gallegos Ruiz Conejo, Ada Lucía

Universidad Nacional Mayor de San Marcos

Datos de Catalogación Bibliográfica

Alarcón Anco, R. J.
Gallegos Ruiz Conejo, A. L.

Inteligencia artificial aplicada a la investigación científica: herramientas, ética y nuevas formas de producir conocimiento

Oriente-Manabí Editorial, Ecuador, 2026
ISBN: 978-9907-9620-3-1
Formato: 210 cm X 270 cm 114 págs.



Publicado por Oriente-Manabí Editorial
Ecuador, Manabí, Cod. Post. 130101.
Contacto: +593 959 723 343
Email: info@omeditorial.com
www.books.omeditorial.com

Director General:	*Dr. Guerrero Bermúdez Ángel Enrique*
Editor en jefe:	*Dr. Guerrero Bermúdez Ángel Enrique*
Editor Académica:	*Lcdo. Oltramonti Roberto, Mg*
Supervisor de Producción:	*Ing. Barragán Monrroy Roberto Johan, Mg.*
Diseño:	*OM Editorial*
Consejo Editorial	*OM Editorial*

© Noviembre, 2026
Libro Digital, Primera Edición, 2026
Editado, Diseñado, Diagramado y Publicado por Comité OM Editorial
Manabí, Ecuador, 2026
D.R. © 2025 por Autores y OM Editorial Ecuador.
Cámara Ecuatoriana del Libro con Radicación editorial 182865

Disponible para su descarga gratuita en www.books.omeditorial.com

Los contenidos de este libro pueden ser descargados, reproducidos difundidos e impresos con fines de estudio, investigación y docencia o para su utilización en productos o servicios no comerciales, siempre que se reconozca adecuadamente a los autores como fuente y titulares de los derechos de propiedad intelectual, sin que ello implique en modo alguno que aprueban las opiniones, productos o servicios resultantes. En el caso de contenidos que indiquen expresamente que proceden de terceros, deberán dirigirse a la fuente original indicada para gestionar los permisos.

Título del libro:
Inteligencia artificial aplicada a la investigación científica: herramientas, ética y nuevas formas de producir conocimiento.

© Alarcón Anco, Ronald Jesús; Gallegos Ruiz Conejo, Ada Lucía.

ISBN: 978-9907-9620-3-1

https://doi.org/10.63618/omeditorial/l43

Como citar (APA 7ma Edición):
Alarcón Anco, R. J., & Gallegos Ruiz Conejo, A. L. (2026). *Inteligencia artificial aplicada a la investigación científica: herramientas, ética y nuevas formas de producir conocimiento*. Oriente-Manabí Editorial. https://doi.org/10.63618/omeditorial/l43

Cada uno de los textos de OM Editorial han sido sometido a un proceso de evaluación por pares doble ciego externos (double-blindpaperreview) con base en la normativa del editorial.

Revisores:

Ing. Juan Manuel Guerrero Calero, PhD	Universidad Estatal del Sur de Manabí – Ecuador
Ing. Víctor Joel Pinargote Bravo, Mgs.	Escuela Superior Politécnica Agropecuaria de Manabí Manuel Félix López – Ecuador

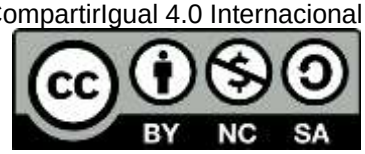
Los libros publicados por **"OM Editorial"** cuentan con varias indexaciones y repositorios internacionales, lo que respalda la calidad de las obras. Lo puede revisar en los siguientes apartados:

Aviso Legal:

La información presentada, así como el contenido, fotografías, gráficos, cuadros, tablas y referencias de este manuscrito es de exclusiva responsabilidad del/los autor/es y no necesariamente reflejan el pensamiento de la OM Editorial.

Derechos de autor ©

Este documento se publica bajo los términos y condiciones de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0).



El “copyright” y todos los derechos de propiedad intelectual y/o industrial sobre el contenido de esta edición son propiedad de la OM Editorial y sus Autores. Se prohíbe rigurosamente, bajo las sanciones en las leyes, la producción o almacenamiento total y/o parcial de esta obra, ni su tratamiento informático de la presente publicación, incluyendo el diseño de la portada, así como la transmisión de la misma de ninguna forma o por cualquier medio, tanto si es electrónico, como químico, mecánico, óptico, de grabación o bien de fotocopia, sin la autorización de los titulares del copyright, salvo cuando se realice confines académicos o científicos y estrictamente no comerciales y gratuitos, debiendo citar en todo caso a la editorial. Las opiniones expresadas en los capítulos son responsabilidad de los autores.

Reseña de Autores

Alarcón Anco, Ronald Jesus

Universidad Nacional Mayor de San Marcos

ronald.alarcona@unmsm.edu.pe

https://orcid.org/0000-0002-7971-5302



Abogado y licenciado en Educación, con maestrías en Docencia Universitaria y Gestión Educativa, así como en Filosofía e Investigación. Doctor en Administración, PhD in Business Administration. Profesor universitario de pregrado y posgrado en la UNMSM y en universidades peruanas e internacionales. Investigador RENACYT-P0051760. Consultor en derechos humanos, defensa nacional e inteligencia artificial aplicada al derecho de la empresa y la educación. Asesor en derecho constitucional, empresarial y educativo. Conferencista en gestión legal, empresarial y sistemas educativos.

Gallegos Ruiz Conejo, Ada Lucía

Universidad Nacional Mayor de San Marcos

agallegosr@unmsm.edu.pe

https://orcid.org/0000-0002-8264-711X



Doctora en Educación y en Gobierno y Políticas Públicas; magíster en Administración Pública. Docente principal de la Universidad Nacional Mayor de San Marcos e investigadora RENACYT. Directora del Instituto de Investigación y Estudios de la Mujer de la Universidad Ricardo Palma. Investiga políticas públicas y gobernanza universitaria y edita revistas científicas indizadas. Becaria Fulbright Humphrey. Fue directora ejecutiva de la Asociación de Universidades del Perú y preside el Consortium for Women Leaders in Public Service. Reconocida por Forbes Perú entre las 50 mujeres más influyentes de 2024.

Índice

Reseña de Autores vii

Índice ix

Índice de Tablas xi

Índice de Figuras xii

Introducción xiii

Capítulo I: La inteligencia artificial y la transformación de la investigación científica 1

1.1. De la automatización a la inteligencia artificial generativa 2

1.2. La colaboración entre investigadores y sistemas inteligentes 6

1.3. Qué significa producir conocimiento científico con IA 9

1.4. Capacidades limitaciones y criterio del investigador 13

1.5. Presentación del caso que acompañará el libro 17

Capítulo II: Del problema de investigación a una revisión de literatura verificable 22

2.1. Delimitación del problema y coherencia del diseño metodológico 23

2.2. Búsqueda y organización crítica de la evidencia 26

2.2.1. Estrategias de búsqueda y selección documentada de estudios 28

2.2.2. Extracción de información y comprobación de referencias 32

2.3. Elección de herramientas según la tarea y los recursos disponibles . 36

2.4. Construcción del protocolo y primeros registros del caso práctico 38

Capítulo III: Análisis de datos y evaluación de los resultados asistidos por IA . 43

3.1. Preparación de los datos y elección del procedimiento analítico 44

3.2. Estrategias de análisis según la naturaleza de los datos 50

3.2.1. Análisis cuantitativo y comprobación de cálculos y código..... 50

3.2.2. Análisis cualitativo y revisión de categorías e interpretaciones... 57

3.3. Contraste de resultados y explicación de errores en el caso práctico 62

Capítulo IV: Ética e integridad a lo largo de la investigación 65

4.1. Responsabilidad del investigador y límites de la delegación 66

4.2. Protección de las personas y los datos 69

4.2.1. Marco peruano del consentimiento y la información sensible 70

4.2.2. Sesgos algorítmicos y representación de las poblaciones 74

4.3. Trazabilidad y revisión de las decisiones 76

4.3.1. Registro de herramientas, versiones, instrucciones y modificaciones 78

4.3.2. Reproducibilidad y evaluación ética del caso práctico 80

Capítulo V: Comunicar hallazgos y producir conocimiento con rigor 85

5.1. Escritura científica y correspondencia entre afirmaciones y evidencia 86

5.2. Autoría y declaración transparente del uso de inteligencia artificial ... 90

5.3. Evaluación del aporte científico y de sus límites 93

5.3.1. Diferencias entre predicción, explicación y evidencia causal 95

5.3.2. Ciencia abierta, colaboración y nuevas formas de descubrimiento 98

5.4. Cierre del caso y criterios transferibles a otras investigaciones 100

Referencias Bibliográficas 105

Índice de Tablas

Tabla 1. Tareas de investigación y responsabilidad del investigador 8
Tabla 2. Ficha inicial del ejemplo didáctico con registros simulados 19
Tabla 3. Propósitos y compromisos de los recorridos documentales 28
Tabla 4. Campos de extracción y condiciones de aceptación de una fuente ... 33
Tabla 5. Protocolo de continuidad del caso con registros simulados 40
Tabla 6. Diccionario operativo y controles de calidad del caso sintético 46
Tabla 7. Descripción y asociación calculadas sobre los sesenta registros sintéticos 52
Tabla 8. Codificación inicial y límites de interpretación de los fragmentos ficticios 59
Tabla 9. Contraste entre productos del ejercicio y procedimientos de comprobación 63
Tabla 10. Matriz de riesgos, responsabilidades y evidencia revisable 69
Tabla 11. Minimización de la entrada: caso utilizado y posible adaptación empírica 74
Tabla 12. Versiones y estados documentados de los productos del caso 79
Tabla 13. Resultados de la auditoría C4-A01 y límites de cada control 82
Tabla 14. Mapa de afirmaciones y evidencia para comunicar el caso simulado 88
Tabla 15. Tareas y controles que puede recoger la declaración de uso de IA . 92
Tabla 16. Procedimientos transferibles y condiciones para su aplicación 103
Tabla 17. Inventario de cierre y estado de los materiales del ejercicio 104

Índice de Figuras

Figura 1. Del planteamiento a una afirmación respaldada 12
Figura 2. Control de incorporación de una fuente 35
Figura 3. Revisión de un análisis asistido por inteligencia artificial 49
Figura 4. Horas de estudio autónomo y puntuación en el archivo sintético 53
Figura 5. Ciclo de trazabilidad de una tarea asistida por IA 77
Figura 6. Controles distintos para el cálculo y la justificación del uso 83
Figura 7. De los resultados a una comunicación verificable 94
Figura 8. Preguntas distintas y evidencia necesaria para responderlas 97

Introducción

La inteligencia artificial está entrando en tareas habituales de la investigación científica, desde organizar una búsqueda bibliográfica hasta proponer código o revisar un manuscrito, y su incorporación obliga a examinar cómo se toman las decisiones que sostienen el conocimiento; este libro aborda ese recorrido desde una perspectiva aplicada, con atención a las herramientas y a las condiciones que permiten utilizarlas con rigor, porque la utilidad de una respuesta depende de la pregunta formulada, de los materiales aportados y de las comprobaciones que el investigador puede realizar antes de aceptarla como parte de su trabajo.

Investigar supone trabajar con información incompleta y elegir entre alternativas que rara vez ofrecen la misma respuesta; una referencia puede parecer pertinente y resultar insuficiente, un dato puede cumplir el formato esperado y conservar un error de unidad, mientras que una explicación bien escrita puede exceder lo observado, de modo que aprender a reconocer esas diferencias resulta necesario para integrar la asistencia tecnológica en un proceso que mantenga sus decisiones abiertas al examen de otras personas.

El propósito de la obra es acompañar a estudiantes de posgrado, docentes y equipos de investigación que necesitan comprender qué pueden delegar y qué deben revisar durante el desarrollo de un estudio; para ello, relaciona la formulación del problema con la búsqueda y lectura de fuentes, avanza hacia el tratamiento de datos y recupera esas decisiones al preparar la comunicación científica, con una secuencia que permite entender por qué un cambio introducido al comienzo puede afectar las conclusiones, aunque sus consecuencias solo se adviertan cuando el manuscrito está casi terminado y corregirlo exige volver a materiales que parecían resueltos.

El contexto latinoamericano atraviesa esta discusión mediante condiciones concretas de acceso, formación y capacidad institucional; elegir una herramienta implica considerar si el equipo puede comprender sus operaciones, conservar los archivos necesarios y continuar el trabajo cuando cambien las condiciones del servicio, además de reconocer que las experiencias y necesidades de una población no quedan representadas automáticamente en los datos disponibles,

por lo que las decisiones metodológicas requieren una lectura atenta del entorno en el que se investiga.

A lo largo del libro, las instrucciones dirigidas a los asistentes se presentan como parte de un procedimiento que debe poder justificarse; su valor reside en delimitar una tarea y facilitar su revisión, con entradas identificables y un estado explícito de las respuestas, para que el lector distinga una propuesta de una operación ejecutada y comprenda qué comprobación hace falta antes de incorporar el resultado a un análisis o a una explicación del fenómeno estudiado.

Un caso didáctico sobre horas de estudio autónomo y puntuación permite seguir el recorrido con materiales que conservan su identidad entre capítulos: sesenta registros sintéticos, acompañados por fragmentos cualitativos ficticios que enseñan otra clase de revisión; la simulación hace posible mostrar errores de preparación, contrastar cálculos y examinar inferencias sin atribuir experiencias a participantes inexistentes, pero también fija un límite que permanece en el cierre, pues una relación incorporada al generador puede servir para comprender el procedimiento y comprobar su funcionamiento, aunque no aporte evidencia sobre el comportamiento de estudiantes reales ni demuestre la eficacia de una herramienta.

El análisis de datos se desarrolla con atención a lo que cada operación permite afirmar, de manera que la descripción de un patrón pueda distinguirse de una predicción y de una explicación causal; esa separación orienta la elección de controles y evita que un coeficiente, una categoría o una representación gráfica reciban un significado que el diseño no sostiene, al tiempo que permite reconocer el aporte de un estudio cuyo objetivo es acotado y cuya utilidad depende de mostrar con claridad lo que efectivamente consiguió.

La ética acompaña estas decisiones desde el inicio y encuentra un desarrollo específico en la protección de las personas, la responsabilidad y la trazabilidad; trabajar con información sensible exige revisar su procedencia y finalidad antes de compartirla, mientras que conservar un registro de instrucciones o versiones ayuda a reconstruir cambios y atender errores, siempre que el equipo explique quién puede examinar los materiales y bajo qué condiciones, porque documentar

una operación no vuelve aceptable cualquier uso de sus resultados ni reemplaza las autorizaciones que correspondan en una investigación con personas.

La comunicación científica reúne lo aprendido durante ese recorrido y lo convierte en afirmaciones que otros puedan valorar; revisar una cita, ajustar una conclusión o declarar el uso de inteligencia artificial requiere conservar la correspondencia con los registros del estudio, así como reconocer las tareas asistidas y la responsabilidad de quienes firman, de modo que la fluidez del texto contribuya a comprender el argumento y permita advertir sus límites sin desplazar la evidencia que lo sostiene.

La lectura puede seguir la secuencia completa o regresar a una decisión particular mediante los ejemplos y recursos de cada capítulo, con la precaución de adaptar los procedimientos al diseño y a la disciplina; el criterio que articula la obra consiste en mantener visible de dónde procede una afirmación, cómo se obtuvo y qué revisión recibió, para que la inteligencia artificial pueda integrarse en una práctica científica que facilite el contraste y la colaboración, conserve sus incertidumbres y ofrezca a otras personas condiciones suficientes para comprender, cuestionar y corregir el conocimiento comunicado.

Capítulo I: La inteligencia artificial y la transformación de la investigación científica



La inteligencia artificial y la transformación de la investigación científica

Investigar con inteligencia artificial exige decidir qué operaciones se delegan, qué información puede utilizar el sistema y mediante qué evidencia se aceptarán sus resultados; esas decisiones afectan la selección de literatura, la construcción de variables, el análisis y la comunicación de los hallazgos, de modo que incorporar una herramienta modifica el trabajo científico cuando cambia alguna de esas operaciones, aunque el documento final conserve la estructura habitual de un artículo o una tesis y aunque la intervención tecnológica parezca limitada a la preparación de un borrador.

Este capítulo desarrolla un criterio de uso basado en la correspondencia entre tarea, evidencia y responsabilidad, con el propósito de que el lector pueda justificar cuándo conviene recurrir a una aplicación y qué debe comprobar antes de utilizar su salida; el recorrido combina estudios experimentales, revisiones y discusiones conceptuales, cuyas contribuciones se distinguen por el tipo de afirmación que permiten sostener, y presenta un ejemplo con registros simulados que hará visibles las decisiones metodológicas a lo largo del libro sin atribuirles resultados de una investigación empírica.

1.1. De la automatización a la inteligencia artificial generativa

La automatización permite ejecutar una operación definida de antemano, como ordenar registros por fecha, identificar duplicados mediante una regla o calcular una media a partir de una columna; su utilidad depende de que las instrucciones representen la tarea y de que los datos cumplan las condiciones necesarias, porque un procedimiento puede funcionar exactamente como fue programado y producir una interpretación equivocada si se aplicó a la variable incorrecta, se confundieron las unidades o se incluyeron observaciones que debían excluirse por razones establecidas en el diseño.

El aprendizaje automático añade la posibilidad de ajustar modelos a partir de datos para reconocer regularidades y utilizarlas en nuevos registros, lo que desplaza parte del trabajo desde la escritura explícita de reglas hacia la elección

de información, objetivos y criterios de evaluación; por ejemplo, estimar una puntuación académica requiere definir qué variables estarán disponibles al predecir y con qué observaciones se evaluará el modelo, mientras que incorporar inadvertidamente información del resultado en los predictores puede ofrecer un rendimiento aparente que desaparezca cuando se intente utilizar el procedimiento en una situación nueva.

En su revisión sobre inteligencia artificial para la ciencia, Wang et al. (2023) describen aplicaciones relacionadas con hipótesis, experimentación y análisis, junto con dificultades de calidad, disponibilidad y gestión de datos, lo que sitúa estas tecnologías dentro de operaciones científicas diversas; el aporte de esa perspectiva consiste en ampliar la discusión más allá de la escritura asistida y en mostrar que la evaluación debe atender al problema concreto, porque detectar una estructura en una imagen, proponer una molécula y organizar literatura requieren evidencias y controles que no pueden reducirse a la calidad de una respuesta conversacional.

La inteligencia artificial generativa incorpora sistemas capaces de producir contenido, como texto o código, a partir de una instrucción y un contexto, pero la forma del producto no informa por sí sola cómo se obtuvo ni qué comprobaciones lo respaldan; una explicación sobre rendimiento académico puede utilizar vocabulario pertinente y una secuencia causal convincente sin haberse contrastado con datos del estudio, mientras que un programa generado puede contener una operación adecuada para otra escala o para otro diseño, por lo que interpretar la salida exige reconstruir la relación entre lo solicitado y lo efectivamente disponible.

La distinción entre predecir y explicar resulta especialmente útil al valorar esas salidas, dado que reconocer asociaciones que anticipan un resultado y establecer qué ocurriría al intervenir sobre una variable plantean exigencias diferentes; Schölkopf et al. (2021) examinan esa relación desde el aprendizaje de representaciones causales y la generalización, ofreciendo una base para comprender por qué una predicción acertada no demuestra por sí misma un mecanismo causal, argumento que en el ejemplo educativo obliga a separar la

estimación de una puntuación de la afirmación de que aumentar horas de estudio produciría una mejora.

También conviene distinguir la generación de contenido de la consulta documental, porque una respuesta apoyada en archivos recuperados introduce materiales que pueden inspeccionarse, aunque todavía sea necesario comprobar cómo se seleccionaron y utilizaron; si el sistema devuelve un párrafo con una referencia, el investigador debe revisar que el documento exista, que el pasaje corresponda al tema y que la interpretación conserve sus condiciones, de manera que disponer de acceso a fuentes amplía las posibilidades de revisión sin convertir automáticamente cada frase en una síntesis correcta o suficientemente representativa de la literatura.

Los beneficios de productividad cuentan con evidencia específica en el experimento de Noy y Zhang (2023), quienes asignaron tareas de escritura profesional a 453 participantes con formación universitaria y encontraron, bajo las condiciones de su estudio, una reducción del tiempo promedio y una mejora de las evaluaciones de los textos; estos resultados respaldan la posibilidad de obtener ganancias en determinadas tareas de redacción, aunque el objeto evaluado no fue un proceso científico completo y sus indicadores no acreditan que se hayan mejorado simultáneamente la validez de los datos, la selección bibliográfica y la justificación de las conclusiones.

El estudio de Gao et al. (2023) aporta un contraste relevante al mostrar que resúmenes científicos generados a partir de títulos y nombres de revistas podían parecer convincentes pese a incorporar información inventada; leído junto con el experimento de Noy y Zhang (2023), permite distinguir dos dimensiones de evaluación, la calidad de una pieza de escritura dentro de una tarea y la correspondencia entre afirmaciones y evidencia, sin presentar ambos trabajos como resultados incompatibles ni extrapolar sus hallazgos a todas las herramientas posteriores, cuyos comportamientos requerirían nuevas evaluaciones.

Cuando una aplicación combina lenguaje, búsqueda y ejecución de programas, la interacción debe describirse mediante las operaciones realizadas, porque proponer un cálculo, ejecutar un archivo y revisar su salida son momentos

diferentes del trabajo; un texto que afirma haber calculado una correlación solo resulta comprobable si pueden identificarse los datos, el procedimiento y el resultado registrado, mientras que un fragmento de código sin ejecución constituye una propuesta técnica que todavía debe probarse, aunque su sintaxis parezca correcta y aunque incluya comentarios que describan con precisión lo que se espera obtener.

Esta diversidad de operaciones permite evaluar una herramienta sin atribuirle una capacidad general a partir de una demostración aislada, pues una aplicación que organiza documentos puede fallar al interpretar una escala y otra que redacta con claridad puede devolver referencias defectuosas; el criterio de selección debe relacionar cada uso con un producto verificable y con el costo de revisarlo, considerando además si una regla convencional o una consulta directa resuelven la tarea con mayor trazabilidad, sin suponer que una operación mejora por el solo hecho de incorporar un componente generativo.

La dimensión institucional amplía ese criterio, ya que el volumen de OECD (2023) examina la IA en la ciencia junto con necesidades de infraestructura, acceso a recursos y colaboración entre especialidades; su enfoque permite argumentar que una adopción sostenible requiere condiciones de trabajo que exceden la adquisición de una aplicación, porque la disponibilidad de datos, el tiempo de revisión y las capacidades disciplinares determinan qué operaciones pueden sostenerse y qué resultados pueden auditarse, aunque instituciones distintas utilicen una interfaz comercial semejante o anuncien políticas de incorporación tecnológica con objetivos similares.

En el recorrido de este libro, estas diferencias se traducirán en una regla práctica: antes de solicitar una salida se definirá si se busca una propuesta, una transformación o una ejecución, y después se conservará la evidencia adecuada a ese producto; tal distinción permitirá aprovechar la velocidad de generación sin convertirla en un indicador de corrección, además de evitar que el relato del procedimiento confunda lo que la herramienta sugirió con lo que el investigador aceptó, modificó o comprobó mediante archivos, documentos originales y resultados que otro lector pueda examinar.

1.2. La colaboración entre investigadores y sistemas inteligentes

La colaboración requiere una distribución de tareas suficientemente precisa para reconocer qué aporta el sistema y qué juicio corresponde a la persona responsable, por lo que una instrucción como “ayuda con la investigación” ofrece menos control que una solicitud para proponer términos de búsqueda o revisar la correspondencia entre variables y objetivos; delimitar la intervención permite comparar alternativas, identificar errores y decidir si la salida merece conservarse, mientras que una delegación demasiado amplia dificulta localizar el punto en que una sugerencia modificó el diseño o introdujo una afirmación sin respaldo.

La presencia de una persona tampoco garantiza que la combinación alcance el mejor desempeño disponible, como muestra el metaanálisis de 106 experimentos de Vaccaro et al. (2024), que distingue mejorar respecto del humano solo de superar al mejor de los dos componentes por separado; las combinaciones evaluadas no alcanzaron esta última condición en promedio y presentaron diferencias según las tareas, de modo que la expresión “supervisión humana” necesita acompañarse de una descripción del arreglo de trabajo y del criterio utilizado para juzgar si esa supervisión contribuye a detectar o corregir errores.

El contraste con Noy y Zhang (2023) ayuda a interpretar ese resultado sin convertirlo en una objeción general a la asistencia tecnológica, porque un estudio de escritura puede documentar mejoras respecto de un grupo de comparación mientras una síntesis de tareas diferentes pregunta si la combinación supera al componente más competente; para un equipo de investigación, la consecuencia metodológica consiste en definir el referente antes de declarar una ganancia, distinguiendo tiempo de elaboración, calidad del producto, errores que permanecen y esfuerzo adicional de revisión, en lugar de reunir todos esos criterios en una valoración única de utilidad.

Las condiciones de revisión también importan en el estudio de Buçinca et al. (2021), quienes compararon intervenciones que exigían mayor participación analítica con formas más simples de presentar recomendaciones y

explicaciones; las primeras redujeron la aceptación de sugerencias incorrectas bajo las condiciones evaluadas, aunque implicaron tensiones con la valoración subjetiva y beneficios desiguales entre participantes, lo que respalda examinar cómo se organiza la decisión y advierte que una interfaz fácil de aceptar no necesariamente promueve el examen más cuidadoso de la recomendación recibida.

Una aplicación editorial de ese argumento consiste en formular primero un criterio propio de selección y solicitar después alternativas al sistema, de manera que el investigador disponga de una razón previa para aceptar o rechazar la sugerencia; si se busca literatura sobre horas de estudio, por ejemplo, conviene definir población, variables y tipo de evidencia antes de consultar resúmenes automáticos, porque ese orden permite detectar que un artículo sobre satisfacción con una plataforma no responde a la misma pregunta que un estudio sobre asociación entre tiempo de estudio y puntuaciones académicas.

La tabla 1 organiza tareas frecuentes mediante cinco componentes que permiten convertir la colaboración en un procedimiento revisable, relacionando el apoyo posible con una comprobación, una decisión y un límite de interpretación; su función es orientar el trabajo del libro y no presentar una clasificación universal de productos, por lo que cada fila deberá adaptarse a las capacidades reales de la herramienta y al diseño utilizado, conservando la posibilidad de rechazar una salida cuando su revisión resulte insuficiente o cuando modifique un supuesto que el investigador no pueda justificar.

Tabla 1.

Tareas de investigación y responsabilidad del investigador

Tarea	Apoyo posible de la IA	Comprobación necesaria	Decisión humana	Límite
Delimitar el problema	Proponer preguntas y variables	Coherencia con el objetivo	Definir qué se estudiará	Una propuesta no es un diseño validado
Revisar literatura	Sugerir términos y organizar registros	Existencia y pertinencia de las fuentes	Seleccionar evidencia y contrastar el original	Una referencia plausible puede ser falsa

Preparar datos	Detectar ausencias y sugerir reglas	Diccionario, rangos y registro de cambios	Aceptar o corregir cada transformación	Depurar puede alterar el sentido de una variable
Analizar información	Proponer código o categorías	Ejecución, supuestos y contraste	Elegir el procedimiento e interpretar	La sugerencia no acredita que se calculó
Comunicar hallazgos	Organizar y revisar un borrador	Correspondencia con resultados y fuentes	Aprobar afirmaciones y límites	La fluidez no demuestra validez

Revisar una respuesta exige dejar evidencia de qué se examinó y qué cambió como consecuencia de ese examen, porque una anotación general de “aprobado” no permite reconstruir si se contrastaron fuentes, unidades o supuestos; en una búsqueda bibliográfica puede registrarse qué términos se conservaron y cuáles se descartaron, mientras que en la preparación de datos conviene preservar el archivo de entrada y las reglas aplicadas, de modo que la responsabilidad humana resulte visible en decisiones concretas y no dependa únicamente de la firma que aparece al final del manuscrito.

En la discusión sobre escritura científica, Salvagno et al. (2023) presentan posibilidades de apoyo para organizar, redactar y revisar materiales, insistiendo en la intervención experta y en la responsabilidad de quienes publican; su aportación corresponde a una perspectiva argumentativa, mientras que Noy y Zhang (2023) ofrecen evidencia experimental sobre tareas de escritura delimitadas, diferencia que permite utilizar ambos trabajos de forma complementaria sin atribuir a la primera una medición de eficacia ni interpretar el segundo como una autorización para transferir al sistema la decisión sobre lo que puede afirmarse científicamente.

La responsabilidad también depende de conservar información suficiente para reconstruir una operación, como plantea el trabajo de Pineau et al. (2021) sobre el programa de reproducibilidad de NeurIPS, que reúne disponibilidad de código, ejercicios de reproducción y listas de comprobación; su experiencia en aprendizaje automático aporta un fundamento para registrar insumos y procedimientos, aunque no constituye una prueba causal de que cualquier lista elimine los errores, por lo que en este libro la documentación se utilizará como

condición que facilita el examen de un resultado y no como sustituto de ese examen.

En una interacción generativa, el registro útil incluirá la instrucción, el contexto suministrado, la identificación disponible de la herramienta, la fecha y las modificaciones aceptadas, reconociendo que una conversación posterior puede producir una respuesta diferente; conservar el intercambio permite estudiar por qué se tomó una decisión aun cuando no pueda obtenerse una reproducción textual exacta, mientras que los cálculos necesitarán además archivos y salidas ejecutadas, porque la trazabilidad de la redacción y la reproducibilidad de una operación estadística responden a preguntas relacionadas que no deben confundirse en el informe metodológico.

La propuesta de van Dis et al. (2023) sitúa la verificación humana y la transparencia entre las prioridades de la investigación con modelos de lenguaje, aportando una orientación que cobra sentido cuando se especifica quién puede revisar cada producto; una corrección de estilo puede resolverse dentro del equipo editorial, pero una decisión sobre muestreo o interpretación necesita competencia metodológica y disciplinar, de modo que conservar la aprobación final resulta insuficiente si la persona responsable carece de información o recursos para examinar la recomendación que se le presenta.

La colaboración será defendible cuando permita identificar un aporte y revisar sus consecuencias, incluso si la decisión final consiste en descartar lo generado, porque detectar una variable improcedente o una referencia irrelevante también mejora el control del proceso; evaluar esa contribución requiere observar el producto completo después de la revisión y no únicamente la rapidez con que apareció una primera respuesta, criterio que guiará el ejemplo conductor al documentar qué propuestas se incorporan al protocolo, cuáles se corrigen y qué razones sostienen cada cambio antes de avanzar hacia el análisis.

1.3. Qué significa producir conocimiento científico con IA

El valor científico de una afirmación depende de la relación que puede establecerse entre su contenido y las razones o evidencias que la respaldan, lo que exige examinar cómo se definió el objeto, mediante qué procedimiento se obtuvo la información y hasta dónde permite interpretarla; una formulación clara

facilita ese examen, pero su apariencia académica no acredita que se hayan cumplido esas condiciones, especialmente cuando el texto combina observaciones, hipótesis y explicaciones sin distinguir qué parte procede del estudio y qué parte corresponde a una posibilidad todavía no contrastada.

El análisis conceptual de Messeri y Crockett (2024) examina promesas de productividad y objetividad asociadas con la IA y describe riesgos de ilusiones de comprensión, en los que la confianza excede lo que se entiende del fenómeno; su contribución permite discutir un problema epistemológico que no se reduce al error factual, dado que un equipo puede utilizar resultados correctos y aun así perder de vista sus supuestos o las alternativas que quedaron fuera del procedimiento, aunque ese argumento no demuestre que toda investigación asistida produzca necesariamente tal efecto.

La discusión de Peralta Palazón (2026) sobre el estatuto epistémico de la IA generativa complementa esa perspectiva al interrogar la relación entre verosimilitud y justificación, desplazando la evaluación desde la apariencia del enunciado hacia las condiciones de su aceptación; leídas conjuntamente, ambas aportaciones permiten formular una exigencia para este libro: cada afirmación relevante debe vincularse con una razón examinable y con un alcance definido, de modo que el lector pueda distinguir una explicación propuesta para orientar el trabajo de una conclusión respaldada por datos o documentos efectivamente revisados.

En la práctica conviene separar tres productos que pueden surgir en una misma interacción: una propuesta sugiere qué investigar o cómo analizar, un resultado informa lo obtenido mediante una operación identificable y una conclusión relaciona ese resultado con la pregunta; esta distinción evita que una hipótesis generada se convierta en hallazgo al redactar el informe y permite exigir controles diferentes, pues una propuesta se evalúa por su pertinencia y posibilidad de contraste, mientras que un resultado necesita evidencia de ejecución y una conclusión requiere examinar la suficiencia de esa evidencia para sostener lo afirmado.

La distinción causal añade una exigencia adicional al mostrar que una asociación observada puede admitir mecanismos diferentes, como señala la revisión de

Schölkopf et al. (2021) al relacionar aprendizaje, intervención y generalización; si dos variables educativas se mueven juntas, esa regularidad todavía no informa qué ocurriría al modificar deliberadamente una de ellas, por lo que una explicación generada debe reconocer qué condiciones exigiría una interpretación causal y qué alternativas permanecen abiertas, en vez de transformar la dirección aparente de una relación estadística en un mecanismo demostrado por la claridad del relato.

En el ejemplo conductor, las horas semanales de estudio representan una dimensión temporal y dejan fuera calidad de atención, estrategias, condiciones de aprendizaje y conocimientos previos, por lo que llamar a esa variable “hábitos de estudio” sin precisar su alcance introduciría una sustitución conceptual; el problema existiría aunque la columna estuviera completa y el cálculo fuera correcto, porque la validez del argumento también depende de que la medida corresponda al concepto invocado, de manera que el control de una intervención de IA debe comenzar antes de recibir una tabla o una interpretación automatizada.

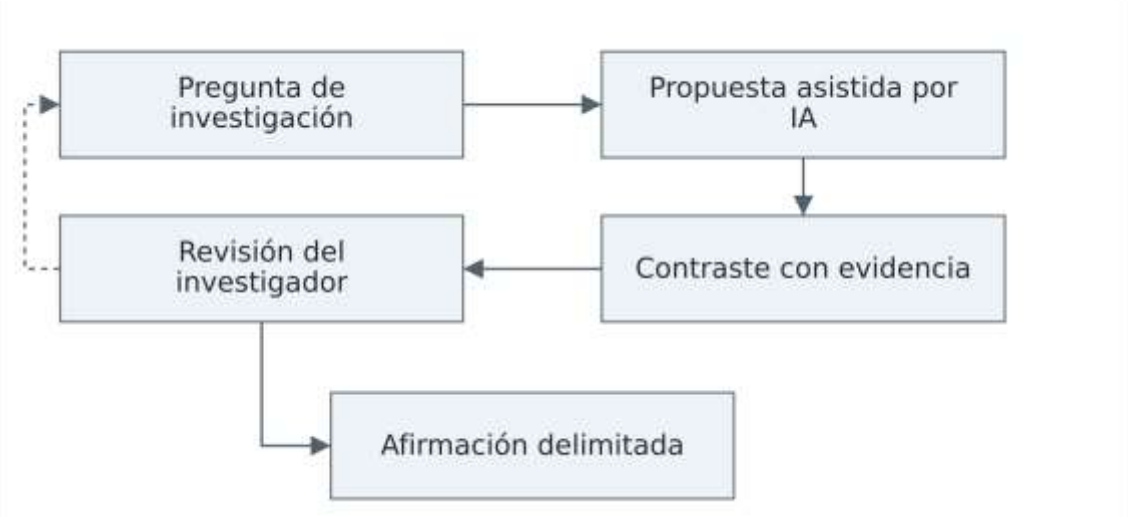
Otra prueba argumentativa consiste en identificar qué información obligaría a revisar una afirmación, porque una explicación que acomoda cualquier resultado difícilmente orienta una comprobación exigente; si se propone que mayor dedicación temporal acompaña puntuaciones superiores, el protocolo debe especificar cómo se observará la relación y cómo se describirán patrones contrarios o débiles, mientras que en un conjunto simulado debe recordarse que las reglas de construcción condicionan lo observado y que examinar esos resultados evalúa el procedimiento sobre un escenario conocido, sin convertirlo en evidencia sobre estudiantes reales.

La figura 1 representa el tránsito desde una pregunta hasta una afirmación delimitada mediante una propuesta, un contraste y una revisión, incluyendo un retorno que permite corregir el planteamiento cuando la evidencia resulta insuficiente; el esquema sirve para evitar una secuencia editorial en la que primero se redacta una conclusión y después se buscan documentos que la acompañen, porque cada paso conserva la posibilidad de modificar o abandonar

la afirmación, condición necesaria para que el uso de una herramienta siga subordinado al problema y a las razones disponibles.

Figura 1.

Del planteamiento a una afirmación respaldada



El prompt C1-P2 propone examinar una afirmación que identifica una explicación convincente con conocimiento científico, con el propósito de hacer visibles los supuestos que sería necesario discutir antes de aceptarla; la salida se presenta como ejemplo elaborado con finalidad didáctica y conserva un alcance conceptual, de manera que el lector pueda valorar el tipo de objeción que orienta la comprobación sin confundirla con una búsqueda documental ejecutada, una evaluación experimental o una prueba independiente de que todas las condiciones científicas han sido satisfechas.

Prompt C1-P2 · Examinar una afirmación

Examina esta afirmación: “Si una IA redacta una explicación convincente, el resultado ya puede considerarse conocimiento científico”. Identifica sus supuestos y describe qué evidencia y controles harían falta. No busques ni inventes referencias; formula una revisión conceptual provisional.

Respuesta ilustrativa para el ejemplo didáctico

La coherencia de una explicación no comprueba su correspondencia con la evidencia; es necesario examinar datos, fuentes, procedimiento y alcance de la inferencia.

Estado del ejemplo: prompt propuesto y respuesta elaborada con finalidad didáctica; no corresponde a una ejecución externa registrada ni a una evaluación experimental de un modelo.

La respuesta ilustrativa identifica una exigencia pertinente al solicitar datos, fuentes y procedimientos, pero todavía carece de materiales concretos con los que realizar ese examen, por lo que su aporte consiste en organizar una comprobación posterior; si el investigador copiara la respuesta como prueba de que la afirmación inicial es falsa, repetiría el salto que pretende corregir al sustituir una formulación convincente por otra, mientras que vincular cada objeción con evidencia identificable permitiría convertir la conversación en un apoyo real para desarrollar el argumento del capítulo.

La variedad de evidencia también debe conservarse cuando se redacta el contraste bibliográfico, dado que un experimento sobre una tarea específica, una revisión de aplicaciones y una discusión epistemológica responden preguntas diferentes; Gao et al. (2023) documentan un comportamiento observado en resúmenes generados, mientras Messeri y Crockett (2024) elaboran un marco para discutir riesgos de comprensión, por lo que relacionar ambos trabajos permite argumentar desde planos complementarios sin tratar la coincidencia temática como una replicación empírica ni atribuir a una revisión conceptual estimaciones que no calculó.

Producir conocimiento con asistencia de IA supone mantener abiertas esas distinciones durante la escritura y el análisis, porque la herramienta puede ampliar alternativas o facilitar una explicación mientras el investigador conserva la obligación de justificar por qué una interpretación merece aceptarse; el aporte científico se reconocerá en la relación que pueda reconstruirse entre pregunta, decisiones, evidencia y conclusión, incluyendo resultados que contradigan las expectativas, y esa relación permitirá valorar la contribución tecnológica de forma más precisa que una declaración general sobre el carácter innovador del procedimiento o la sofisticación de la aplicación utilizada.

1.4. Capacidades limitaciones y criterio del investigador

La selección de una herramienta debe comenzar por una operación concreta y por sus requisitos de revisión, considerando qué error sería relevante para el argumento y qué evidencia permitiría detectarlo antes de incorporar la salida;

una aplicación puede ser adecuada para organizar archivos y resultar insuficiente para interpretar un diseño, mientras que generar código exige comprobar su ejecución y los supuestos del análisis, por lo que atribuir confiabilidad a la herramienta en conjunto oculta diferencias entre tareas y dificulta asignar controles proporcionales a las consecuencias de una equivocación.

La bibliografía ofrece un ejemplo directo de esa exigencia en el estudio de Walters y Wilder (2023), quienes examinaron referencias generadas por las versiones de ChatGPT que evaluaron y documentaron tanto publicaciones fabricadas como errores en registros correspondientes a obras existentes; la diferencia resulta decisiva para revisar un manuscrito, porque encontrar un título real resuelve únicamente una parte del problema y todavía deja pendiente comprobar autores, año, identificador y pertinencia, sin utilizar las tasas observadas en aquel estudio como si describieran de manera permanente todas las versiones y servicios posteriores.

El contraste con Gao et al. (2023) amplía la discusión desde la identidad de una referencia hasta la consistencia de un texto científico, dado que una apariencia convincente puede coexistir con información inventada dentro del resumen; ambos trabajos permiten argumentar que la revisión debe seguir el contenido hacia su evidencia, aunque estudian objetos y procedimientos distintos, por lo que no corresponde reunir sus resultados como si ofrecieran una única tasa de error, sino utilizarlos para reconocer puntos diferentes del proceso en los que una presentación formal puede ocultar problemas de respaldo.

Para este libro se aplican tres controles complementarios a cada fuente utilizada: comprobar su identidad mediante el registro editorial y un identificador o catálogo pertinente, localizar en el original el respaldo de la afirmación y revisar después la correspondencia entre ese pasaje y la redacción final; estos controles forman una pauta de trabajo editorial y no equivalen a declarar que tres personas independientes revisaron cada cita, ya que su valor proviene de las preguntas diferentes que resuelven y de la posibilidad de conservar evidencia de cada comprobación realizada.

El primer control permite detectar un título inexistente, una autoría incorrecta o un identificador que conduce a otra obra, pero no demuestra que la publicación respalde la interpretación utilizada; el segundo exige leer resultados, método y contexto cuando sean necesarios para comprender el pasaje, mientras que el tercero examina si el capítulo convirtió una posibilidad en certeza, una asociación en causalidad o un resultado de una tarea en una conclusión general, además de comprobar versiones y correcciones relevantes antes de cerrar la referencia y su hipervínculo.

Los detectores de texto generado tampoco resuelven esa revisión de contenido, como muestran desde ángulos distintos Liang et al. (2023) y Weber-Wulff et al. (2023), cuyos estudios examinan, respectivamente, sesgos en muestras de escritura en inglés y limitaciones de herramientas de detección; sus resultados corresponden a corpus y sistemas determinados, pero ofrecen razones para evitar que una puntuación de detector sustituya el examen de fuentes, datos y argumentos, puesto que identificar un probable origen textual y evaluar la corrección científica del contenido constituyen tareas diferentes.

La redacción natural se trabajará mediante precisión, variedad sintáctica y continuidad argumentativa, conservando las decisiones y atribuciones que permiten revisar el texto, en lugar de convertir una calificación automática de estilo en objetivo de calidad; una frase puede requerir reescritura porque confunde categorías o exagera un hallazgo, aunque ningún detector la marque, mientras que una formulación convencional puede ser adecuada si expresa con exactitud el procedimiento, por lo que la revisión editorial debe justificar los cambios por su contribución al significado y a la claridad para el lector.

El perfil de gestión de riesgos de Autio et al. (2024) reúne riesgos y acciones sugeridas para la IA generativa, incluyendo problemas de información incorrecta y condiciones de uso, dentro de un marco voluntario que debe adaptarse al contexto; su utilidad para la investigación consiste en organizar preguntas sobre insumos, controles y consecuencias, sin presentarlo como una norma obligatoria para una institución ecuatoriana, porque la evaluación de una aplicación necesita distinguir una orientación técnica de las obligaciones efectivamente aplicables y de las autorizaciones que exige el entorno institucional donde se trabaja.

La protección de información requiere examinar qué materiales se transferirán y bajo qué condiciones, como plantea la orientación institucional de UNESCO (2026) para educación e investigación, cuya página actualizada remite a una guía publicada originalmente en 2023; esta distinción de fechas evita presentar la actualización del sitio como una nueva edición del informe y permite utilizar su orientación sobre capacidades humanas y datos con una atribución precisa, mientras que cada estudio deberá establecer qué información puede compartirse antes de incorporar archivos o conversaciones a un servicio externo.

El contexto regional aporta otra dimensión en el libro de Jácome-Jácome et al. (2026), que aborda infraestructura, conectividad y brechas de acceso en América Latina, en diálogo con las condiciones de integración examinadas por OECD (2023); ambas perspectivas permiten sostener que la factibilidad depende también de recursos y organización, sin presumir que las instituciones comparten un diagnóstico idéntico, por lo que una estrategia viable puede requerir copias autorizadas, continuidad del trabajo y formación situada antes de ampliar las operaciones que se encomiendan a una herramienta.

La literatura educativa permite precisar el alcance de esa formación mediante las oportunidades y desafíos discutidos por Kasneci et al. (2023) y la revisión documental de Jácome-Jácome (2026) sobre personalización en educación superior; su lectura conjunta aporta preguntas sobre mediación, transparencia y preparación docente, aunque no demuestra una mejora causal uniforme en cualquier intervención, de modo que el ejemplo del libro utilizará ese contexto para justificar decisiones y límites, conservando la diferencia entre una discusión sobre posibilidades pedagógicas y una evaluación empírica de resultados de aprendizaje.

El criterio del investigador se expresa al distinguir errores corregibles de restricciones del diseño, porque una referencia mal transcrita debe repararse mientras el alcance limitado de una muestra o de una variable debe mantenerse en la interpretación; también requiere reconocer cuándo una tarea necesita otra competencia profesional o una fuente adicional, de manera que la capacidad para redactar instrucciones se integre con conocimientos disciplinares, metodológicos y editoriales, permitiendo aceptar una ayuda útil sin atribuirle

autoridad sobre decisiones cuyas razones todavía no se han examinado suficientemente.

1.5. Presentación del caso que acompañará el libro

El caso conductor examinará la relación entre horas de estudio autónomo y una puntuación académica mediante registros simulados de estudiantes universitarios adultos, con el propósito de mostrar cómo se toman y revisan decisiones asistidas por IA; no corresponde a una universidad ni a participantes reales, por lo que su valor didáctico residirá en hacer observables las operaciones y sus límites, permitiendo que el lector siga la misma pregunta a través de búsqueda, preparación de datos, análisis y escritura sin confundir el desarrollo del ejemplo con evidencia sobre una población estudiantil.

La expresión “hábitos de estudio” reúne prácticas heterogéneas que no pueden representarse mediante una sola cifra sin precisar qué dimensión se selecciona, por lo que el caso comenzará con el tiempo semanal de estudio autónomo, excluyendo horas de clase; esta delimitación facilita construir un diccionario de datos y revisar las transformaciones propuestas, aunque deja fuera organización, estrategias y calidad de atención, de modo que cualquier redacción posterior deberá conservar la diferencia entre la variable temporal utilizada y el concepto más amplio que motivó el tema de investigación.

La pregunta operativa será qué asociación se observa entre las horas semanales de estudio autónomo y una puntuación académica dentro del conjunto de registros simulados, formulación que identifica variables y restringe el alcance al archivo construido; esa restricción evita prometer una explicación causal o una estimación poblacional que el ejemplo no puede proporcionar, además de ofrecer un criterio para revisar las respuestas de la herramienta cuando proponga objetivos adicionales, cambie la población o sugiera que una relación en datos sintéticos demuestra una mejora real en el aprendizaje de estudiantes.

El diseño prevé 60 registros con códigos S001 a S060, valores de tiempo entre 0 y 40 horas por semana y una puntuación entre 0 y 100, límites elegidos para construir un escenario manejable y examinar su consistencia; estas cifras no proceden de una encuesta, no expresan una recomendación sobre dedicación académica y no representan parámetros estimados, por lo que deben

mantenerse identificadas como decisiones de simulación tanto en el archivo y las tablas como en los prompts que se utilicen para generar o interpretar el material.

El número de registros facilitará la inspección del conjunto y el seguimiento de operaciones sin presentarse como un cálculo de potencia o de representatividad, ya que aumentar el tamaño de un archivo sintético no lo convierte en una muestra de personas reales; antes de generarlo se documentarán reglas y semilla, conservando una versión estable para las comprobaciones posteriores, en consonancia con el valor de registrar datos y procedimientos que discuten Pineau et al. (2021), aunque la reproducibilidad técnica del ejemplo tampoco acreditará que sus supuestos describan adecuadamente una población universitaria.

Las reglas de generación deberán permitir reconstruir por qué aparecen determinados patrones, incluyendo el componente aleatorio y cualquier relación incorporada deliberadamente entre variables, porque una asociación introducida en el diseño del archivo no constituye un descubrimiento sobre comportamiento académico; esa transparencia permitirá utilizar el conjunto para examinar la ejecución de un procedimiento y reconocer errores de interpretación, mientras que ocultar las reglas y presentar el patrón resultante como evidencia educativa confundiría el funcionamiento de una simulación con la observación de un fenómeno que nunca fue medido en participantes reales.

La tabla 2 fija la identidad del caso mediante componentes, decisiones y condiciones de interpretación, de manera que las etapas posteriores puedan revisarse contra una referencia estable y no contra descripciones que cambien según la respuesta de la herramienta; si se decide agregar una variable o modificar una escala, el cambio deberá registrarse y justificarse antes de utilizarlo, conservando las versiones necesarias para reconocer qué análisis corresponde a cada archivo y evitando que una modificación inadvertida altere el significado de los resultados que se comparan o comunican.

Tabla 2.

Ficha inicial del ejemplo didáctico con registros simulados

Componente	Decisión de construcción	Alcance o condición
Pregunta	Asociación entre horas de estudio autónomo y puntuación académica	Limitada al conjunto sintético
Unidad de análisis	Un registro sintético de estudiante universitario adulto	No identifica a una persona ni a una institución real
Tamaño previsto	60 registros, con códigos S001 a S060	Elegido para seguimiento didáctico; no es una muestra empírica
Tiempo de estudio	Horas por semana, entre 0 y 40; excluye horas de clase	Dimensión temporal de los hábitos de estudio
Puntuación académica	Escala construida entre 0 y 100	No equivale a notas observadas en una universidad
Supuesto operativo	Archivo único y reglas de generación documentadas antes de analizar	La asociación puede depender de las reglas elegidas
Inferencia permitida	Descripción y asociación dentro del archivo sintético	No demuestra causalidad, representatividad ni mejora real

El prompt C1-P1 solicita apoyo para delimitar una pregunta de asociación, definir una unidad y proponer variables, estableciendo que el ejemplo es simulado y que no deben inventarse citas ni atribuirse datos a instituciones; su interés metodológico reside en especificar la tarea y sus restricciones, aunque ninguna instrucción garantiza por sí misma que la respuesta las respete, por lo que la salida se acompaña de una revisión editorial que permitirá reconocer qué decisiones se aceptan, cuáles necesitan precisión y qué evidencia faltaría antes de presentar el diseño como un estudio ejecutado.

Prompt C1-P1 · Delimitar la pregunta

Actúa como apoyo metodológico. Quiero construir un ejemplo didáctico, con datos simulados, sobre hábitos de estudio y rendimiento académico en estudiantes universitarios adultos. Propón una pregunta de asociación, define la unidad de análisis y sugiere variables medibles. Distingue las decisiones que necesitan criterio del investigador. No atribuyas datos a personas o instituciones reales, no inventes citas y no presentes el diseño como ejecutado.

Respuesta ilustrativa para el ejemplo didáctico

¿Qué asociación se observa entre las horas semanales de estudio y una puntuación académica en el conjunto de datos simulados?; la unidad de análisis propuesta es un registro sintético de estudiante, y antes de analizar deberán fijarse la escala del puntaje y la definición de hora de estudio.

Estado del ejemplo: prompt propuesto y respuesta elaborada con finalidad didáctica; no corresponde a una ejecución externa registrada ni a una evaluación experimental de un modelo.

La respuesta ilustrativa propone una pregunta compatible con el propósito y una unidad sintética, pero todavía requiere precisar que las horas corresponden a estudio autónomo y que la puntuación utiliza una escala construida, decisiones que se incorporan a la ficha; este examen muestra que una sugerencia razonable puede ser incompleta sin resultar enteramente inútil y que revisar consiste en identificar su contribución y corregir su alcance, dejando constancia de los cambios en lugar de describir la aceptación del investigador como una validación por expertos o como evidencia aportada por participantes.

La relación entre tiempo y puntuación también permitirá discutir explicaciones alternativas, como conocimientos previos, dificultad de las asignaturas o condiciones de dedicación, que en un estudio empírico exigirían decisiones de medición y diseño; la distinción entre asociación e intervención examinada por Schölkopf et al. (2021) ayuda a mostrar por qué esos problemas no desaparecen al obtener una predicción ajustada, mientras que en el caso sintético tales alternativas se utilizarán para delimitar la interpretación y no para fingir que se han controlado factores cuya información no forma parte del archivo construido.

Durante el recorrido se conservarán la pregunta aprobada, el diccionario, el archivo y las salidas de los procedimientos que realmente se ejecuten, junto con las instrucciones y correcciones relevantes de la interacción, para poder examinar situaciones como una etiqueta confundida con una medida o una conclusión que exceda los registros; estos materiales permitirán distinguir una propuesta de análisis de un cálculo realizado y una ilustración editorial de un resultado obtenido, evitando que la continuidad narrativa del caso borre las diferencias entre lo planificado, lo ejecutado y lo que todavía requiere comprobación.

El paso al capítulo siguiente consistirá en relacionar pregunta, búsqueda documental y protocolo, estableciendo qué evidencia sería pertinente para contextualizar el problema y qué decisiones deben quedar fijadas antes de analizar el archivo; los criterios transferibles a otras investigaciones serán la claridad de las variables, la trazabilidad de los cambios y la correspondencia entre resultados y afirmaciones, mientras que tamaño, escalas y reglas de simulación permanecerán sujetos al propósito didáctico de este caso, de modo que trasladar el procedimiento exigirá adaptar y justificar esas decisiones frente al nuevo problema.



Capítulo II: Del problema de investigación a una revisión de literatura verificable

Del problema de investigación a una revisión de literatura verificable

Una investigación comienza a adquirir forma cuando el interés general se convierte en una pregunta que puede responderse con información identificable, dentro de un alcance que el diseño permite defender; la inteligencia artificial puede ayudar a explorar formulaciones, reconocer vocabulario y organizar documentos, pero esas operaciones solo contribuyen al estudio cuando mantienen una relación explícita con el problema y permiten recuperar las razones de cada decisión, porque una colección abundante de textos o una pregunta bien redactada todavía pueden conducir a un trabajo cuyo objetivo, variables y evidencia no coincidan.

El capítulo desarrolla el tránsito entre la delimitación del problema y la preparación de un protocolo, integrando búsqueda, selección, extracción y elección de herramientas como decisiones conectadas; el caso sobre horas de estudio autónomo y puntuación académica avanza hacia una ficha metodológica utilizable, mientras las cadenas de búsqueda y las respuestas de ejemplo conservan su condición de propuestas didácticas, de modo que el lector pueda seguir un procedimiento concreto sin atribuirle recuentos documentales, participantes o ejecuciones que no corresponden a esta etapa del recorrido.

2.1. Delimitación del problema y coherencia del diseño metodológico

Delimitar un problema exige establecer qué relación merece examinarse, sobre qué unidades y mediante qué información, evitando que el tema general ocupe el lugar de una pregunta investigable; “inteligencia artificial y educación” puede orientar una exploración inicial, pero reúne intervenciones, poblaciones y resultados demasiado diferentes para determinar un análisis, mientras que preguntar por una asociación entre dos variables dentro de un conjunto definido permite decidir qué documentos son pertinentes, qué medidas deben precisarse y qué afirmaciones quedarían fuera del estudio, aun cuando parezcan relacionadas con el mismo campo temático.

La justificación debe explicar por qué conocer esa relación aporta información y qué decisión o comprensión podría mejorar, conservando una distancia entre la relevancia social del tema y la contribución específica del trabajo; un problema educativo importante puede requerir un estudio modesto y bien delimitado antes que un diseño amplio que no pueda ejecutarse, por lo que conviene describir el vacío mediante preguntas y diferencias entre evidencias disponibles, evitando declarar que “no existen investigaciones” cuando únicamente se ha realizado una búsqueda preliminar o se dispone de acceso parcial a ciertas fuentes.

En el tratamiento de la formación investigativa de Torres-Vela et al. (2025), la operacionalización y la elección de estrategias cuantitativas aparecen vinculadas con la preparación metodológica del investigador; esta orientación permite situar el uso de IA dentro de una competencia que exige comprender qué se mide y para qué se analiza, aunque la discusión del libro no valida por sí misma un instrumento particular ni garantiza la adecuación de cada propuesta tecnológica, por lo que aquí se traduce en una exigencia concreta de coherencia entre pregunta, variable y procedimiento previsto.

Los objetivos deben formular operaciones que puedan reconocerse en los resultados, de manera que describir una distribución, estimar una asociación y evaluar una intervención no se utilicen como expresiones intercambiables; si el archivo contiene horas semanales y puntuaciones, un objetivo descriptivo puede organizar ambas variables y otro relacional examinar su comportamiento conjunto, mientras que afirmar que se evaluará una mejora causada por estudiar más exigiría decisiones adicionales sobre intervención, comparación y temporalidad, elementos que no aparecen por el simple hecho de pedir a un asistente que utilice un lenguaje más técnico.

La unidad de análisis determina qué representa cada fila y qué comparaciones pueden realizarse, por lo que una base con varios registros por estudiante necesita un tratamiento distinto de otra que contiene una sola observación por persona; también debe distinguirse la unidad sobre la que se recopila información de aquella a la que se pretende atribuir una conclusión, porque describir registros de una asignatura no autoriza automáticamente afirmaciones sobre toda una institución y porque una salida agregada puede ocultar

diferencias que resultan relevantes para interpretar la pregunta originalmente planteada.

El libro de Velasco-Holguin et al. (2026) examina el rendimiento universitario como una construcción multidimensional y sitúa sus indicadores dentro de trayectorias y condiciones educativas, argumento que ayuda a precisar el significado de la puntuación del caso; utilizar una escala entre cero y cien ofrece una variable manejable para la demostración, pero no representa todas las dimensiones del aprendizaje ni permite atribuir características personales a quienes obtendrían un valor determinado, de modo que el nombre de la columna y la interpretación deberán conservar ese alcance desde la preparación del diccionario.

Una definición operacional útil debe indicar qué se registrará, en qué unidad, durante qué período y con qué reglas de inclusión, de modo que dos personas puedan reconocer si están trabajando con la misma variable; en el caso, las horas corresponderán al estudio autónomo semanal y excluirán el tiempo de clase, mientras que la puntuación será una escala construida para el ejemplo, decisión que impide tratarla como una calificación observada y permite detectar propuestas de IA que mezclen horas diarias, dedicación mensual o medidas de satisfacción bajo una etiqueta común.

La temporalidad también debe establecerse antes de analizar, porque un total semanal y una puntuación acumulada durante un semestre no representan necesariamente períodos comparables; cuando la información procede de un estudio real, esa diferencia exige justificar el vínculo entre las medidas o revisar el diseño de recolección, mientras que en una simulación corresponde explicitar la convención adoptada y conservarla durante el recorrido, evitando que una explicación posterior atribuya a una variable un significado temporal que el archivo nunca incluyó o que la instrucción original dejó sin definir.

El metaanálisis de Aeon et al. (2021) encuentra relaciones entre gestión del tiempo, desempeño y bienestar, pero su objeto incluye prácticas de organización que exceden la cantidad de horas dedicadas a una actividad; esta diferencia conceptual permite contextualizar el interés educativo del ejemplo sin utilizar la revisión como si proporcionara parámetros para generar sus registros, ya que

una asociación entre gestión temporal y logro no equivale a demostrar una relación idéntica entre horas de estudio autónomo y puntuación académica ni autoriza a trasladar sus estimaciones a un archivo sintético construido con otro propósito.

La asistencia generativa puede ser útil para proponer objetivos alternativos o advertir contradicciones entre términos, siempre que el investigador conserve una versión aprobada del planteamiento y registre las modificaciones que cambien su sentido; si una respuesta transforma “asociación” en “impacto”, el cambio debe examinarse como una alteración del alcance inferencial y no como una mejora estilística, porque aceptar esa palabra podría comprometer al capítulo de resultados con una explicación causal que los datos y el diseño no pueden sostener, aunque el resto de la redacción parezca más convincente.

La revisión inicial del diseño debe comprobar si cada objetivo necesita una variable disponible y si cada análisis propuesto contribuye a responder la pregunta, incluyendo la posibilidad de retirar operaciones que añadan complejidad sin aportar información; esta comprobación evita construir un procedimiento alrededor de las funciones de una herramienta y permite reconocer carencias antes de generar datos o seleccionar documentos, momento en que todavía es posible ajustar el alcance con razones explícitas y mantener un estudio realizable, coherente con sus recursos y suficientemente preciso para orientar la búsqueda bibliográfica.

2.2. Búsqueda y organización crítica de la evidencia

La revisión de literatura construye el contexto argumental del estudio mediante documentos seleccionados con un propósito reconocible, por lo que su calidad depende de la relación entre búsqueda, criterios y uso de la información; una revisión narrativa puede organizar debates y conceptos, una sistemática responder una pregunta delimitada mediante procedimientos explícitos y una revisión de alcance cartografiar tipos de evidencia o vacíos, mientras que una búsqueda exploratoria ayuda a comprender el vocabulario y ajustar el planteamiento, funciones que conviene distinguir antes de elegir una plantilla o solicitar una síntesis automática.

La declaración de Page et al. (2021) presenta PRISMA 2020 como una guía para informar revisiones sistemáticas mediante una lista de veintisiete elementos y recursos complementarios, cuyo propósito es hacer comprensible qué se hizo y qué se encontró; esa función permite utilizarla para mejorar el reporte sin convertir el cumplimiento formal de una lista en prueba de exhaustividad o calidad metodológica, por lo que un diagrama bien presentado deberá corresponder a decisiones y registros reales, mientras la adecuación del diseño y la evaluación de los estudios requieren criterios específicos.

La tabla 3 compara propósitos y compromisos de distintos recorridos documentales para facilitar una elección proporcional al trabajo, sin establecer una jerarquía automática en la que toda investigación deba transformarse en revisión sistemática; la elección debe responder a la pregunta y a la contribución esperada, porque contextualizar un caso simulado y sintetizar efectos de intervenciones educativas plantean necesidades diferentes, y porque prometer un procedimiento que no se podrá sostener puede generar más opacidad que declarar con precisión el alcance y las decisiones de una revisión narrativa documentada.

Tabla 3.

Propósitos y compromisos de los recorridos documentales

Recorrido	Propósito	Procedimiento que debe explicitarse	Alcance del producto
Búsqueda exploratoria	Reconocer vocabulario y ajustar la pregunta	Recursos consultados, términos y cambios relevantes	Orientación inicial; no representa un censo
Revisión narrativa documentada	Organizar conceptos y contrastar argumentos	Criterios de pertinencia, rutas de búsqueda y fuentes utilizadas	Síntesis interpretativa con alcance declarado
Revisión sistemática	Responder una pregunta delimitada mediante una síntesis planificada	Protocolo, búsqueda, selección, evaluación y síntesis según diseño	Conclusión condicionada por evidencia y método

Revisión de alcance	Describir tipos de evidencia, conceptos y vacíos de un campo	Protocolo, criterios y procedimientos de selección y organización	Mapa de evidencia; no implica estimar un efecto combinado

La orientación de Peters et al. (2022) para desarrollar protocolos de revisiones de alcance detalla la preparación previa de la pregunta, los criterios y las etapas de trabajo, situando el protocolo como una guía que permite anticipar cómo se organizará la evidencia; su aporte resulta pertinente cuando el objetivo consiste en describir un campo y sus características, aunque no convierte cualquier colección de publicaciones en una revisión de alcance, de modo que adoptar esa denominación implica justificar su propósito y mantener coherencia entre el mapa que se pretende producir y el procedimiento diseñado.

En el caso conductor, la literatura cumplirá una función de fundamentación y precisión conceptual del protocolo, por lo que se preparará una revisión narrativa documentada con una estrategia exploratoria explícita, sin anunciar un metaanálisis ni un censo exhaustivo de estudios; esta decisión permite combinar fuentes metodológicas y educativas según su contribución, conservando criterios de pertinencia y trazabilidad, mientras que un proyecto futuro orientado a estimar asociaciones en una población real tendría que definir un recorrido de selección y síntesis acorde con esa pregunta y con la evidencia disponible.

2.2.1. Estrategias de búsqueda y selección documentada de estudios

Una estrategia comienza por identificar los conceptos indispensables y sus expresiones alternativas, antes de unir términos con operadores o introducir filtros que puedan reducir la recuperación; para el caso, población universitaria, dedicación al estudio y desempeño académico ofrecen bloques iniciales que deberán revisarse mediante documentos pertinentes, porque el vocabulario cotidiano, los descriptores de una base y las expresiones utilizadas por autores de distintas disciplinas pueden referirse a objetos próximos sin coincidir exactamente en la forma con la que se nombran las variables dentro del protocolo.

La traducción entre español e inglés requiere conservar el significado del concepto y reconocer sus límites, dado que “study time”, “time management” y “study habits” pueden conducir a literaturas relacionadas con medidas diferentes; mantenerlos como alternativas exploratorias puede ayudar a localizar documentos, pero el criterio de inclusión tendrá que distinguir qué práctica se estudió realmente, evitando que la expansión terminológica se convierta después en una equiparación de resultados que medían duración, organización o estrategias de aprendizaje bajo definiciones que no permiten compararlos directamente.

La extensión PRISMA-S de Rethlefsen et al. (2021) propone dieciséis elementos para informar búsquedas bibliográficas, incluyendo fuentes, estrategias y procedimientos que permiten comprender cómo se identificó la evidencia; su utilidad consiste en transformar una frase general sobre “búsqueda en bases científicas” en un registro examinable, sin imponer idéntica sintaxis a todos los servicios, por lo que el protocolo deberá conservar las cadenas efectivamente utilizadas y las condiciones de cada consulta, en lugar de reconstruirlas de memoria después de conocer los documentos que se decidió citar.

La combinación booleana debe reflejar una decisión sobre relaciones conceptuales, usando alternativas dentro de un bloque y conexiones entre bloques cuando cada componente resulte necesario para la recuperación; añadir todos los términos del tema mediante AND puede excluir documentos relevantes, mientras una secuencia extensa de OR sin agrupación puede recuperar materiales ajenos al problema, de modo que las pruebas iniciales deben examinar si aparecen registros conocidos y si los resultados responden al sentido de la pregunta, conservando las variantes que se compararon y la razón de los cambios introducidos.

El análisis de Gusenbauer (2022) compara la cobertura disciplinar de cincuenta y seis bases bibliográficas y muestra diferencias que impiden tratar sus contenidos como equivalentes, aportando una razón para seleccionar recursos según el campo y el propósito de búsqueda; ese hallazgo no establece un orden permanente de calidad entre servicios ni asegura que una combinación cubra todo tema educativo, por lo que la elección del caso considerará especialización,

idiomas y funciones de consulta, evitando identificar el número total de resultados con la suficiencia de la evidencia pertinente para el problema.

Una biblioteca especializada en educación, un índice multidisciplinar y un catálogo editorial pueden cumplir funciones complementarias, siempre que se indique qué se espera encontrar en cada recurso y cómo se integrará después la información; los catálogos de libros permiten localizar desarrollos conceptuales y metodológicos, mientras que las revistas ofrecen estudios o debates con objetos más delimitados, por lo que conviene conservar el tipo documental y la procedencia desde el ingreso del registro, en vez de presentar todos los materiales como si hubieran sido recuperados mediante una misma consulta y evaluados con idénticos criterios.

La guía de la National Library of Medicine (2026) explica que la interpretación de frases, campos y operadores modifica la búsqueda en PubMed, ejemplo que hace visible la necesidad de comprobar la sintaxis en la documentación del servicio utilizado; las etiquetas que funcionan en esa plataforma no deben trasladarse automáticamente a un recurso educativo o a un catálogo editorial, de manera que el uso de IA para proponer una cadena deberá acompañarse de una adaptación verificable de campos y filtros antes de describirla como una consulta efectivamente realizada.

El prompt C2-P1 organiza una propuesta inicial de vocabulario y cadena booleana para que el lector pueda identificar qué parte corresponde a cada concepto, manteniendo la diferencia entre formular una consulta y ejecutarla; su salida ilustrativa requiere depuración semántica y adaptación al recurso elegido, pero permite mostrar una intervención concreta de IA que puede revisarse sin inventar artículos o recuentos, por lo que el ejemplo conserva la cadena como texto copiable y deja la fecha efectiva, los filtros y el archivo exportado para el registro de una ejecución posterior.

Prompt C2-P1 · Proponer una estrategia exploratoria

A partir de la pregunta sobre asociación entre horas de estudio autónomo y puntuación académica en universitarios adultos, propone términos en español e inglés y una cadena booleana inicial. Separa población, dedicación y desempeño; distingue horas, hábitos y

gestión del tiempo. Indica qué sintaxis deberá comprobarse en cada recurso. No ejecutes búsquedas ni inventes artículos, recuentos o DOI.

Respuesta ilustrativa para el ejemplo didáctico

Cadena exploratoria propuesta: ("university students" OR "undergraduate students") AND ("study time" OR "study hours") AND ("academic performance" OR "academic achievement"); en español pueden explorarse "estudiantes universitarios", "tiempo de estudio" y "rendimiento académico", manteniendo "time management" y "study habits" como expansiones que requieren revisar su correspondencia conceptual.

Estado del ejemplo: prompt propuesto y respuesta elaborada con finalidad didáctica; no corresponde a una ejecución externa registrada ni a una evaluación experimental de un modelo.

El rastreo de referencias aporta una vía complementaria en el trabajo de Haddaway, Grainger y Gray (2022), que presenta Citationchaser para explorar referencias citadas y publicaciones citantes a partir de un conjunto inicial; este procedimiento puede localizar literatura que utiliza vocabulario distinto del previsto en la cadena, aunque su cobertura depende de los registros y conexiones disponibles, por lo que una fuente encontrada mediante esa ruta debe conservar su procedencia y someterse a los criterios del estudio, sin tratar la cercanía dentro de una red de citas como prueba suficiente de pertinencia o calidad.

La selección debe separar la pertinencia temática de la suficiencia de información y de la evaluación metodológica, porque un documento puede responder al tema y, aun así, no ofrecer los datos necesarios para el uso previsto; durante una primera lectura se pueden conservar registros potencialmente pertinentes para examinar el texto completo, mientras que la decisión final debe indicar una razón reconocible de incorporación o exclusión, evitando categorías generales como “no aporta” cuando el problema concreto consiste en otra población, una medida incompatible o un diseño que no responde a la pregunta.

Los duplicados requieren un control que combine identificadores y datos bibliográficos, conservando el registro que permita acceder al original y documentando las decisiones ambiguas, porque una misma publicación puede aparecer con pequeñas variaciones de título o autoría; también debe distinguirse la duplicación del registro de la existencia de varios informes de un mismo

estudio, situación en la que eliminar documentos por semejanza podría retirar información complementaria y mantener una única fila por archivo podría contar varias veces la evidencia de los mismos participantes dentro de una síntesis posterior.

La estructura propuesta por Haddaway, Rethlefsen, et al. (2022) distingue información sobre plataforma, base, fechas, estrategias y condiciones de consulta para favorecer búsquedas transparentes y repetibles, aportando un modelo de organización que puede adaptarse al tamaño del proyecto; para este libro se utilizará un registro por ejecución con la cadena exacta y el archivo exportado, manteniendo las propuestas en un estado separado, de modo que una consulta corregida no reemplace silenciosamente la anterior ni un valor vacío de registros recuperados se interprete como una búsqueda sin resultados.

2.2.2. Extracción de información y comprobación de referencias

La extracción transforma el texto de las publicaciones en información comparable mediante campos definidos, por lo que necesita un formulario relacionado con el argumento que se quiere construir y un localizador que permita regresar al original; registrar únicamente una frase de resultados dificulta valorar si el hallazgo procede de una asociación, de una comparación o de una intervención, mientras que incluir diseño, población, medida y límites facilita decidir qué puede incorporarse al capítulo y qué información debe conservarse como contexto sin utilizarla como demostración de una afirmación más amplia.

Antes de ampliar la matriz conviene probar sus campos con documentos de características distintas y revisar dónde aparecen ambigüedades, como la diferencia entre el número de personas reclutadas, quienes completaron una medición y quienes entraron al análisis; una columna única de “muestra” puede ocultar esas distinciones y producir cifras contradictorias al redactar, por lo que el formulario debe permitir identificar a qué etapa corresponde cada número y qué fragmento lo respalda, manteniendo separados el dato extraído, la interpretación del investigador y las dudas que necesitan una nueva lectura.

La prueba de concepto de Gartlehner et al. (2024) obtuvo una exactitud del 96,3 % con Claude 2 al extraer ciento sesenta elementos de diez publicaciones de ensayos, mientras Mahmoudi et al. (2025) encontraron en su última iteración un

93,3 % para características explícitas y un 50 % para componentes conductuales más interpretativos; la diferencia entre tareas y corpus impide comparar esas cifras como una clasificación universal de modelos, pero respalda distinguir campos literales y juicios contextuales al decidir qué extracción puede asistirse y qué revisión exige el contenido antes de incorporarlo al argumento.

La tabla 4 organiza campos mínimos y criterios de aceptación que permiten examinar la extracción sin convertirla en una transcripción extensa del artículo, preservando el documento, el localizador y el uso previsto como partes de un mismo registro; el formulario puede ampliarse cuando una pregunta requiera información adicional, pero cada columna debe responder a una decisión de análisis o interpretación, porque recopilar datos que nunca se utilizarán aumenta el trabajo y puede dar una apariencia de precisión que no mejora la correspondencia entre las fuentes seleccionadas y las afirmaciones del capítulo.

Tabla 4.

Campos de extracción y condiciones de aceptación de una fuente

Campo	Información que se conserva	Criterio de aceptación
Identidad y V1	Autoría, año, título, DOI o ISBN y ficha de origen	Los datos corresponden al documento consultado
Diseño y unidad	Tipo de estudio, población o material y unidad analizada	La descripción procede del original y conserva su alcance
Medida y resultado	Variable, escala, cifra o argumento pertinente y denominador	Se distingue un dato informado de una interpretación
Localizador y V2	Página, sección, tabla o párrafo que respalda el campo	El pasaje puede recuperarse y sostiene la afirmación
Versión y V3	Versión utilizada, fecha de comprobación y cambios editoriales	El uso final corresponde a la versión y al contenido revisados
Uso y decisión	Argumento que respaldará y razón de incorporación	La pertinencia se justifica respecto de la pregunta

El prompt C2-P2 solicita trabajar exclusivamente con un documento aportado y obliga a asociar cada campo con una ubicación, además de permitir declarar que un dato no está informado cuando esa ausencia se haya comprobado; su utilidad depende de que el texto suministrado sea legible y suficiente para la tarea, por lo que una extracción desde un resumen o un fragmento no debe describirse

como lectura del artículo completo, y una salida sin localizador necesita revisión antes de utilizarse aunque coincida con lo que el investigador esperaba encontrar.

Prompt C2-P2 · Extraer campos con localizador

Utiliza exclusivamente este fragmento didáctico simulado D-C2-01, párrafo 1: "Se propone examinar la asociación entre horas semanales de estudio autónomo y una puntuación de 0 a 100 en 60 registros sintéticos de estudiantes universitarios adultos, con una observación por código". Extrae propósito, unidad, número de registros, variables e instrumento de medición; añade el localizador. Si el fragmento no informa un campo, decláralo y no lo completes con conocimiento externo.

Respuesta ilustrativa para el ejemplo didáctico

Propósito: examinar una asociación; unidad: registro sintético con una observación por código; número previsto: 60; variables: horas semanales de estudio autónomo y puntuación de 0 a 100; localizador de estos campos: D-C2-01, párrafo 1; instrumento de medición: no informado en el fragmento aportado, que describe una simulación y no una recolección empírica.

Estado del ejemplo: prompt propuesto y respuesta elaborada con finalidad didáctica; no corresponde a una ejecución externa registrada ni a una evaluación experimental de un modelo.

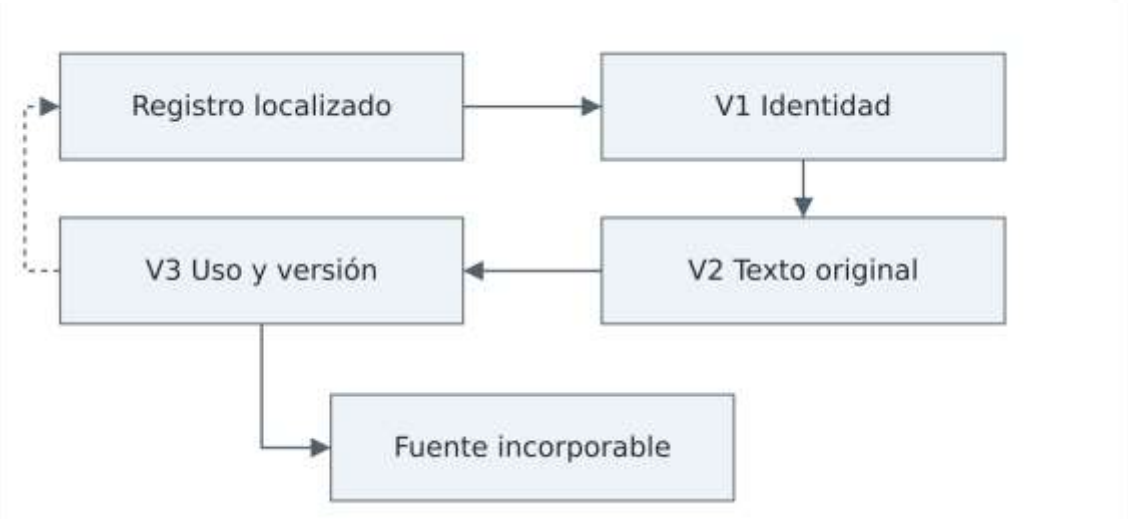
Cuando dos pasajes parecen presentar cifras distintas, la comprobación debe examinar si describen momentos, grupos o denominadores diferentes antes de elegir uno de ellos, preservando la notación y las condiciones con las que se informó el resultado; una proporción calculada sobre respuestas válidas no puede trasladarse a todas las personas invitadas sin justificar el cambio, del mismo modo que un intervalo o una medida de asociación no deben separarse de las variables y del procedimiento que les dan sentido, aunque el asistente ofrezca una reformulación aparentemente más sencilla.

El control bibliográfico comienza con la identidad de la obra y continúa con el contenido y la versión utilizada, de modo que la triple verificación se convierte en una secuencia de preguntas diferentes: si la publicación corresponde al registro, si sostiene la afirmación y si el uso final conserva ese respaldo; la figura 2 representa el paso entre esos controles y la incorporación de la fuente, incluyendo un retorno cuando se necesita corregir información, sin presentar el

esquema como un diagrama de selección de estudios ni como evidencia de recuentos de una revisión ejecutada.

Figura 2.

Control de incorporación de una fuente



La actualización del registro científico requiere una consulta específica, como explica Rittman (2025) al describir la incorporación de información de Retraction Watch a la API de Crossref, donde los metadatos pueden ayudar a localizar retractaciones y otros cambios; ese recurso complementa la ficha editorial y los avisos asociados a la publicación, pero la ausencia de una alerta en una sola consulta no garantiza que nunca exista una modificación, por lo que el control debe conservar fecha, identificador y versión examinada, distinguiendo una corrección puntual de una retirada que altere la posibilidad de utilizar la fuente.

La redacción de una cita debe reunir el aporte pertinente, sus condiciones y la interpretación necesaria en el mismo párrafo, avanzando después hacia otra decisión o evidencia que haga progresar el argumento; así se evita dedicar varios párrafos consecutivos a reformular una misma publicación y se conserva una relación clara entre fuente y afirmación, mientras los desacuerdos entre estudios se resolverán examinando medidas, diseños y contextos dentro de una comparación explícita, sin forzar coincidencias ni aumentar artificialmente el respaldo mediante varias menciones del mismo trabajo.

2.3. Elección de herramientas según la tarea y los recursos disponibles

Elegir herramientas requiere identificar el problema operativo que debe resolverse y el producto que permitirá revisar su cumplimiento, porque organizar referencias, priorizar registros y extraer datos son tareas distintas aunque puedan aparecer dentro de una misma plataforma; una decisión razonable puede combinar un gestor bibliográfico, una matriz y un asistente de lenguaje cuando cada componente tenga una función definida, conservando una ruta de trabajo que el equipo pueda sostener y evitando incorporar aplicaciones cuyos resultados no puedan exportarse, examinarse o relacionarse con los documentos utilizados como entrada.

El costo relevante incluye preparación, revisión y mantenimiento, además de las condiciones de acceso que correspondan al servicio, por lo que una respuesta rápida puede implicar un esfuerzo considerable si devuelve campos difíciles de comprobar o requiere corregir repetidamente los mismos errores; antes de ampliar su uso conviene probar una tarea delimitada con materiales cuya respuesta pueda revisarse, registrar qué se obtuvo y comparar el trabajo completo con la alternativa disponible, sin convertir una demostración favorable en una afirmación de superioridad para todas las etapas de la investigación.

El marco de ASReview presentado por van de Schoot et al. (2021) utiliza aprendizaje activo para priorizar títulos y resúmenes a partir de decisiones de revisión, mostrando mediante simulaciones posibilidades de reducir trabajo de cribado; su aporte corresponde a la organización de un conjunto de registros ya disponible y necesita entenderse dentro de las condiciones evaluadas, por lo que usar un orden de prioridad exige mantener criterios y decisiones verificables, además de justificar cómo se cerrará la revisión sin convertir los registros que quedan al final de la lista en exclusiones automáticas.

La evaluación del cribado debe prestar atención a los estudios pertinentes que podrían perderse, porque una alta proporción global de decisiones correctas puede coexistir con la exclusión de una parte importante de los registros relevantes cuando abundan los documentos ajenos al tema; una prueba útil debe observar por separado inclusiones y exclusiones, además de conservar los

casos dudosos para revisión, de manera que el criterio de éxito responda a las consecuencias de omitir evidencia y no únicamente al número de decisiones que la herramienta consigue resolver con rapidez.

El estudio de Khraisha et al. (2024) examinó GPT-4 en cribado y extracción de literatura en varios idiomas y mostró que interpretar su desempeño cambia al considerar el desequilibrio de los conjuntos y el acuerdo por azar; sus resultados impiden valorar una automatización mediante exactitud aislada y obligan a precisar tarea, criterio y documentos evaluados, sin extrapolar el comportamiento observado a todas las versiones posteriores, por lo que el protocolo del caso reservará la decisión de incorporación al investigador y tratará las clasificaciones automáticas como apoyos que todavía deben examinarse.

La compatibilidad de archivos puede determinar la utilidad práctica de un servicio, especialmente cuando el material incluye tablas, documentos escaneados o diseños que dificultan recuperar el texto en el orden correcto; la prueba inicial debe comprobar si se conservan encabezados, unidades y correspondencias entre columnas, porque una extracción que omite una nota o mezcla dos grupos puede alterar el significado de una cifra aun cuando el texto resultante parezca legible, de modo que el acceso a un PDF no debe confundirse con una representación fiel de toda la información que contiene.

La revisión de software para analítica del aprendizaje de Afzaal y Nouri (2024) diferencia aplicaciones según sus propósitos y características, ofreciendo un ejemplo de evaluación orientada a tareas dentro de la educación superior; ese enfoque resulta útil para evitar que la presencia de un panel, una predicción o una función automatizada se interprete como evidencia de la misma capacidad, aunque una revisión de programas educativos no determine qué gestor bibliográfico necesita este caso, por lo que aquí se retiene el criterio de comparar funciones pertinentes y condiciones de uso antes de adoptar una herramienta.

La continuidad del trabajo requiere que los registros importantes permanezcan disponibles fuera de una conversación o de una sesión concreta, con nombres y versiones que permitan identificar documento, extracción y decisión, respetando las condiciones de acceso a los materiales; cuando la conectividad o las cuentas institucionales sean limitadas, pueden organizarse consultas por etapas y

conservarse exportaciones autorizadas para la revisión posterior, de modo que el protocolo no dependa de recordar una pantalla ni de repetir una interacción para recuperar la única copia de una decisión metodológica relevante.

En Valencia-Galarza et al. (2026), la gobernanza universitaria de la IA incluye responsabilidades, selección de herramientas y transparencia sobre datos y supervisión, lo que sitúa la adopción dentro de una organización concreta; esta perspectiva permite justificar que bibliotecas, equipos técnicos y responsables académicos intervengan cuando la elección excede una tarea personal, sin asumir que toda institución necesita la misma estructura, por lo que el caso distinguirá el uso de documentos públicos de cualquier escenario futuro que requiera acceso a información de estudiantes o autorizaciones adicionales para procesarla.

La selección quedará razonada cuando pueda explicarse qué operación resuelve cada componente, qué información recibe, qué salida conserva y quién examina el resultado, incluyendo una alternativa para continuar si cambia el acceso; este criterio permite utilizar soluciones sencillas cuando satisfacen el propósito y reservar aplicaciones más complejas para necesidades justificadas, mientras las especificaciones temporales deberán comprobarse en la documentación vigente al ejecutar la tarea, evitando que el protocolo dependa de una lista de marcas o de funciones que podrían variar durante la elaboración del libro.

2.4. Construcción del protocolo y primeros registros del caso práctico

El protocolo del caso convierte las decisiones anteriores en un acuerdo de trabajo que puede revisarse antes del análisis, con una pregunta de asociación y dos objetivos operativos: describir las variables y examinar su relación dentro del archivo sintético; la unidad será un registro simulado de estudiante universitario adulto y el diseño tendrá una finalidad didáctica de análisis transversal, sin reclutamiento ni observación de personas, de manera que la preparación del documento establezca qué se podrá ejecutar y qué conclusiones permanecerán fuera del alcance aunque una herramienta las proponga durante la interacción.

Se conservarán sesenta códigos, de S001 a S060, horas semanales de estudio autónomo entre cero y cuarenta y una puntuación entre cero y cien, manteniendo las definiciones aprobadas para facilitar la continuidad del recorrido; el diccionario añadirá el tipo de dato, la unidad y las reglas de validación de cada campo, además de especificar cómo se representaría una ausencia si se incorporara deliberadamente al escenario, evitando que el valor cero se utilice indistintamente para indicar ausencia de estudio, falta de información o una observación que se decidió retirar del análisis.

En la discusión de Lakens (2022), la justificación del tamaño de muestra debe relacionarse con el objetivo inferencial y con la información que se espera obtener, distinguiendo argumentos de potencia, precisión o restricciones de recursos; aplicado por contraste al ejemplo, este criterio obliga a explicar que los sesenta registros responden a la inspección didáctica del procedimiento y no a una estimación de suficiencia para una población real, porque la repetición técnica de un archivo construido no reemplaza la justificación que necesitaría una investigación empírica antes de recopilar datos.

La generación del archivo se realizará en la etapa de preparación de datos, documentando reglas, componente aleatorio y semilla cuando corresponda, mientras este capítulo fija las condiciones que esa generación deberá respetar; cualquier dependencia introducida entre horas y puntuación quedará declarada como una decisión del escenario, permitiendo que el análisis posterior compruebe el comportamiento del procedimiento bajo condiciones conocidas sin presentar el patrón resultante como un hallazgo educativo, y conservando una versión estable para que las comparaciones entre salidas no mezclen archivos distintos bajo el mismo nombre de caso.

La tabla 5 reúne decisiones del protocolo que requieren continuidad entre capítulos y distingue el estado documental del estado de ejecución, de manera que una definición aprobada no aparezca como una actividad ya realizada; su uso permitirá revisar cambios propuestos por una herramienta y reconocer cuándo afectan la pregunta, la escala o el alcance del análisis, además de mantener una referencia común al redactar tablas y conclusiones, evitando que una descripción metodológica aparentemente más completa atribuya al ejemplo

procedimientos de recolección, validaciones o resultados que no se han producido.

Tabla 5.

Protocolo de continuidad del caso con registros simulados

Componente	Decisión fijada	Estado o límite
Pregunta	Asociación entre horas de estudio autónomo y puntuación académica	Limitada al archivo sintético
Objetivos	Describir ambas variables y examinar su relación	No evaluar el efecto de una intervención
Unidad y tamaño	60 registros con códigos S001 a S060	Diseño didáctico; sin participantes reales
Medidas	Horas semanales de 0 a 40 y puntuación de 0 a 100	La puntuación es construida; no son notas observadas
Búsqueda C2-B01	Bloques y cadena propuesta; adaptación al recurso elegido	Ejecución y recuentos se registrarán cuando se realicen
Generación y análisis	Reglas, semilla, validación y archivo estable	Planificados para la etapa siguiente; sin resultados calculados aquí
Versiones	Registrar qué cambia, por qué y desde cuándo	Conservar el vínculo entre protocolo y archivo utilizado

El primer registro de búsqueda podrá identificarse como C2-B01 y contendrá la pregunta vigente, los bloques conceptuales, la cadena propuesta y el recurso al que se adaptará, con los campos de ejecución diferenciados hasta que exista una consulta conservada; cuando se realice, se incorporarán fecha, filtros, cantidad recuperada y nombre del archivo exportado, sin reemplazar la cadena inicial ni asignar un cero a los campos todavía no ejecutados, porque ese cero expresaría un resultado documental que podría interpretarse erróneamente como ausencia de literatura sobre el tema.

La biblioteca del caso organizará los documentos según su función en la investigación, distinguiendo las fuentes que ayudan a definir las variables de las que orientan búsqueda, extracción o análisis, sin imponer que cada publicación responda directamente a la pregunta asociativa; cada registro conservará autoría, año, identificador y acceso al original, mientras una matriz separada recogerá el contenido pertinente y su uso, permitiendo revisar por qué se citó

una obra y evitando que almacenar un archivo se interprete como evidencia de que ya fue leído, evaluado y aceptado para cualquier afirmación del capítulo.

La revisión de Aucapiña-Tubón et al. (2026), publicada en Innova Science Journal, analiza veintiocho artículos y organiza problemas de equidad, protección de datos y transparencia en el contexto educativo ecuatoriano, aportando una perspectiva situada para anticipar decisiones institucionales; su contribución se utiliza como contexto documental y no como prueba de una relación entre horas y puntuación, de modo que el protocolo podrá reconocer condiciones educativas relevantes sin trasladar sus conclusiones a las variables del caso ni presentar la síntesis de esos autores como un estudio empírico ejecutado dentro de este libro.

El plan de análisis comenzará por comprobar códigos, rangos y coherencia de las unidades, seguido de una descripción de ambas variables y una inspección gráfica de su relación, dejando la elección del coeficiente y sus condiciones para el examen del archivo generado; esta secuencia evita decidir una técnica únicamente porque aparece en una respuesta de IA y permite justificarla según los datos y la pregunta, manteniendo la diferencia entre una operación planificada y un resultado calculado que deberá acompañarse del archivo, el procedimiento y la salida correspondiente.

Las modificaciones al protocolo se registrarán mediante una versión, el componente afectado y una razón que explique su necesidad, conservando la distinción entre corregir una definición ambigua y cambiar el objetivo después de observar un patrón; si se agrega una variable para desarrollar un ejemplo complementario, deberá indicarse qué pregunta responde y cómo modifica la construcción del escenario, porque ampliar el archivo de forma silenciosa podría producir una continuidad narrativa aparente mientras las tablas y conclusiones dejan de referirse al mismo conjunto de decisiones metodológicas que orientó el caso.

Al cerrar esta etapa, el producto utilizable será un protocolo definido, una propuesta de búsqueda identificable y una matriz con criterios de extracción, junto con las convenciones que permiten registrar las ejecuciones posteriores; la calidad de esa preparación se reconocerá en la correspondencia entre pregunta,

medidas y decisiones, de modo que el capítulo siguiente pueda avanzar hacia la preparación y el análisis de datos conservando la trazabilidad del recorrido, mientras la revisión documental seguirá abierta a nueva evidencia pertinente que justifique cambios explícitos en lugar de añadirse únicamente para aumentar el número de referencias.

Capítulo III: Análisis de datos y evaluación de los resultados asistidos por IA



Análisis de datos y evaluación de los resultados asistidos por IA

La utilidad de un análisis asistido por inteligencia artificial se decide cuando sus resultados pueden contrastarse con los datos y con la pregunta que los originó; una tabla completa, un programa que termina sin errores o una interpretación bien escrita todavía necesitan esa comprobación, porque cada producto puede conservar decisiones equivocadas que la presentación vuelve difíciles de advertir, desde una unidad confundida hasta una conclusión causal construida sobre una asociación.

En este capítulo se ejecuta el caso de los sesenta registros sintéticos definido anteriormente y se examina cómo cambia su lectura al preparar la base, elegir el procedimiento y comprobar las operaciones; el recorrido incluye una demostración cualitativa complementaria y dos instrucciones de trabajo, con distinción explícita entre los resultados calculados, las propuestas para interactuar con un asistente y los ejemplos de respuesta elaborados para enseñar a detectar errores.

3.1. Preparación de los datos y elección del procedimiento analítico

El archivo que recibe un investigador contiene valores, pero también decisiones sobre qué se midió, cómo se registró y qué se dejó fuera; antes de pedir un análisis conviene recuperar ese significado, pues dos columnas llamadas «tiempo» y «rendimiento» no bastan para saber si se comparan horas semanales con una prueba, minutos diarios con una calificación acumulada o medidas cuya temporalidad impide responder la pregunta prevista, aun cuando el sistema reconozca ambas como números y proponga inmediatamente una correlación; la documentación de conjuntos de datos propuesta por Gebru et al. (2021) incluye motivación, composición, obtención, preparación y usos recomendados, ofreciendo información para juzgar la adecuación de una base antes de utilizarla; trasladar ese criterio al trabajo cotidiano exige registrar qué representa cada fila y qué transformaciones ha recibido, sin atribuir a una ficha documental la

capacidad de corregir por sí misma los sesgos o las carencias del material descrito.

Una revisión inicial puede ser breve cuando la base es sencilla, aunque debe resolver las ambigüedades que afectarían el análisis; en el ejemplo conductor, cada fila representa un registro construido de estudiante universitario adulto, las horas corresponden al estudio autónomo durante una semana convencional y la puntuación pertenece a una escala de cero a cien, por lo que no se dispone de información sobre estrategias de aprendizaje, asistencia, conocimientos previos o condiciones familiares que permita incorporarlos después como explicaciones comprobadas; si en una investigación real la puntuación cambiara de escala entre grupos, la equivalencia tendría que justificarse antes de reunir los valores; una operación aritmética de conversión puede conservar los límites numéricos y, aun así, dejar sin resolver diferencias entre instrumentos, contenidos o condiciones de evaluación.

El identificador merece un tratamiento separado: S001 y S060 permiten reconocer registros y comprobar duplicaciones, mientras sus cifras carecen de significado cuantitativo para la pregunta; convertirlas en un predictor introduciría un orden administrativo como si describiera una característica educativa, error que puede pasar inadvertido si se entregan todas las columnas a una rutina automática sin especificar su función.

La tabla 6 recoge el diccionario operativo y los controles aplicados al caso; además de fijar unidades y rangos, distingue los campos de entrada de las variables que entran efectivamente en el análisis, separación necesaria para conservar la procedencia de una conversión sin mezclarla con las medidas que se comparan.

Tabla 6.

Diccionario operativo y controles de calidad del caso sintético

Campo	Definición y escala	Unidad y rango	Control aplicado
id	Identificador nominal del registro	S001–S060	Sesenta códigos únicos; no entra en la correlación

tiempo_registrado	Valor de entrada antes de normalizar	Horas o minutos, según unidad	Conservar valor y unidad de procedencia
unidad	Etiqueta nominal de la unidad de entrada	horas / minutos	Dividir entre 60 solo cuando consta minutos
horas_semana	Estudio autónomo semanal; variable cuantitativa	0–40 h/semana; una decimal	Rango y ausencia; excluye tiempo de clase
puntuacion	Puntuación cuantitativa construida	0–100 puntos; una decimal	Rango y ausencia; recuperación documentada desde origen sintético

Hay una diferencia práctica entre detectar una inconsistencia y conocer su reparación: un valor de 1830 en una columna de horas semanales viola el rango previsto, pero ese dato aislado no informa si hubo una conversión omitida, un desplazamiento de columna o un error de digitación; en este ejercicio se conserva el indicador «minutos» y la base maestra sintética, de modo que puede justificarse la operación 1830 ÷ 60 = 30,5, mientras que sin ese respaldo correspondería revisar la procedencia antes de reemplazar el valor por otro que pareciera razonable.

En el tratamiento técnico de datos, McKinney (2022) desarrolla operaciones de detección de ausencias, sustitución de valores y eliminación de duplicados, mostrando cómo implementarlas de manera explícita; la decisión metodológica comienza precisamente donde termina la instrucción informática, porque disponer de una función para rellenar celdas permite efectuar la transformación, pero no establece qué valor representa de forma defendible una observación desconocida ni qué efectos tendrá sobre la relación estudiada.

El cero debe conservar su significado: puede expresar que no se dedicaron horas al estudio durante el período definido, mientras una celda vacía indica que esa información no está disponible; sustituir una ausencia por cero cambia una condición de conocimiento por una afirmación sobre la conducta, con consecuencias para medias, dispersiones y asociaciones que no quedan resueltas al declarar que el archivo ya no contiene celdas vacías; también faltan datos cuando una fila completa nunca llega a la base; Wickham et al. (2023) distinguen ausencias explícitas e implícitas y explican cómo las transformaciones

de estructura pueden hacerlas visibles u ocultarlas, distinción que obliga a revisar el conjunto de unidades esperadas y no únicamente el porcentaje de celdas incompletas, porque un archivo sin valores ausentes puede haber perdido registros durante una unión, una exportación o un filtro previo.

La estrategia para tratar información incompleta depende del objetivo y del proceso que produjo la ausencia; Lee et al. (2021) proponen en el marco TARMOS planificar ese tratamiento, examinar su adecuación y comunicar las decisiones junto con sus consecuencias, incluyendo análisis de sensibilidad cuando proceda, de modo que eliminar registros completos o imputar no aparezcan como soluciones universales basadas únicamente en alcanzar una tabla aparentemente ordenada; la razón de la ausencia también puede estar relacionada con el resultado: quienes encuentran mayores dificultades podrían omitir precisamente la información que interesa analizar; sin conocer ese proceso, conservar solo las respuestas disponibles puede producir un conjunto internamente completo que represente de manera desigual las situaciones del estudio.

En la demostración se introdujeron cuatro incidencias conocidas después de generar la base maestra: una repetición exacta de S012, el registro de S009 en minutos, una puntuación vacía en S023 y el valor 105 en S037; la entrada pasó a tener 61 filas, aunque seguía representando sesenta identificadores, por lo que el primer recuento útil consistió en diferenciar filas de unidades y conservar ambos valores en el registro de depuración; las reparaciones se fijaron antes de calcular la asociación: retirar la copia exacta, convertir la unidad documentada y recuperar de la base maestra las puntuaciones 62,9 y 80,8 de S023 y S037; la recuperación fue posible porque el ejercicio conserva su origen conocido, condición que no debe trasladarse a un estudio empírico como si las respuestas perdidas pudieran reconstruirse con la misma certeza, y tampoco equivale a haber imputado valores mediante un modelo.

Un dato extremo válido requiere otra decisión; si una persona hubiera declarado una dedicación inusual pero posible, su distancia respecto del resto no bastaría para eliminarla, porque retirar valores con la intención de obtener una relación más nítida modifica la pregunta después de observar la respuesta; corresponde

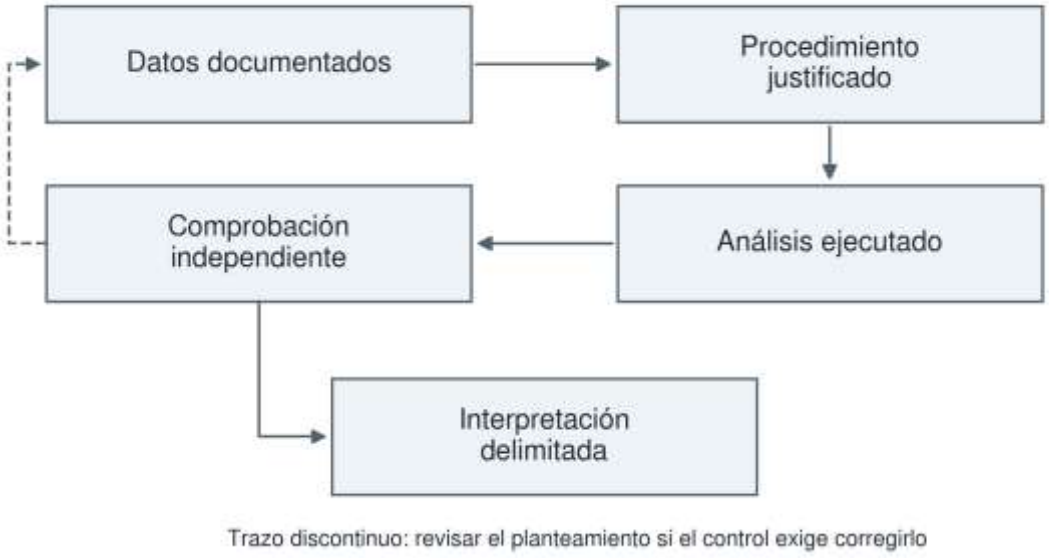
examinar el registro y, cuando sea pertinente, comparar análisis que permitan reconocer su influencia sin convertir la incomodidad del dato en prueba de error; para separar ambas situaciones resulta útil comprobar el documento de origen antes de consultar qué sucede con el coeficiente al excluir el registro; ese orden conserva una razón de calidad anterior al beneficio estadístico y deja abierta la posibilidad de mantener un valor incómodo cuando su procedencia y significado sean consistentes.

La calidad de la preparación incluye mantener una secuencia reconocible entre entrada, transformación y salida; en este caso se conservaron la base maestra, la copia con incidencias y la base analítica, con un registro que identifica la fila afectada y la regla aplicada, lo que permite revisar por qué el análisis utiliza sesenta pares completos y evita que la versión depurada sustituya silenciosamente el material que hizo necesaria la corrección; la trazabilidad debe abarcar también las acciones pequeñas: cambiar un separador decimal o normalizar una etiqueta puede modificar qué filas reconoce una condición de filtrado, por lo que conviene registrar la regla aplicada al conjunto y conservar ejemplos suficientes para examinar si actuó de la manera prevista.

La figura 3 muestra el recorrido y su retorno posible: si la comprobación revela una discrepancia, habrá que revisar la preparación o el procedimiento antes de interpretar.

Figura 3.

Revisión de un análisis asistido por inteligencia artificial



Elegir el procedimiento exige volver a los objetivos: describir las horas, examinar su asociación con la puntuación y predecir un resultado futuro son tareas diferentes, aunque una misma biblioteca permita ejecutarlas consecutivamente; la primera necesita resúmenes adecuados a la distribución, la segunda una medida coherente con el tipo de relación y la tercera una estrategia de evaluación fuera del ajuste, de manera que añadir operaciones disponibles en la interfaz puede ampliar el informe sin mejorar la respuesta al problema; el diseño puede imponer restricciones adicionales, como mediciones repetidas o unidades agrupadas que invaliden la suposición de independencia; esos requisitos deben discutirse antes de que el asistente elija una prueba por reconocer únicamente el formato de las columnas, porque una hoja de cálculo no describe por sí sola cómo se produjeron sus filas.

La asistencia será más útil si recibe ese marco y debe explicar la correspondencia entre objetivo, escala y operación; solicitar «el análisis más avanzado» deja abierta una selección orientada por complejidad, mientras pedir una propuesta justificada sobre dos variables definidas permite detectar una clasificación innecesaria, una transformación que destruye información o una interpretación que ya supone una intervención inexistente.

3.2. Estrategias de análisis según la naturaleza de los datos

Una base numérica y un conjunto de relatos requieren formas distintas de comprobación; en la primera pueden verificarse sumas, unidades y resultados de funciones, mientras en los segundos interesa recuperar los pasajes que sostienen una interpretación y reconocer qué decisiones del analista los reúnen bajo una categoría, por lo que la calidad de ambos recorridos no puede juzgarse mediante un único porcentaje de coincidencia con la respuesta del asistente.

La diferencia tampoco establece una jerarquía entre contar e interpretar; una media puede estar bien calculada y describir una variable mal construida, del mismo modo que una categoría convincente puede estar desconectada de los fragmentos que supuestamente la sostienen, de manera que la revisión debe examinar tanto la operación efectuada como la pertinencia de aquello que se eligió operar; la selección de evidencias debería comenzar por aquello que se desea conocer: si interesa comprender por qué una persona organiza el estudio de cierta manera, una cifra de dedicación semanal aporta un contexto distinto del que ofrece su relato sobre interrupciones, acuerdos familiares o dificultad para concentrarse.

El libro de James et al. (2023) organiza el aprendizaje estadístico distinguiendo objetivos de predicción e inferencia y desarrollando procedimientos para ajustar y evaluar modelos; esa distinción ayuda a fijar el lugar del caso, cuya asociación se calcula dentro de un archivo sintético y no constituye una evaluación de capacidad predictiva para nuevos estudiantes, aun cuando sea posible representar los mismos pares mediante una recta ajustada.

3.2.1. Análisis cuantitativo y comprobación de cálculos y código

La descripción comienza por observar qué valores aparecen y cómo se distribuyen; promediar antes de verificar el rango puede ocultar una unidad incorrecta, mientras informar únicamente una media válida deja fuera diferencias que afectan su lectura, razón por la cual el caso presenta tamaño, media, desviación estándar muestral, mediana y extremos observados, manteniendo separadas las unidades de horas por semana y puntos de la escala construida.

Las reglas de simulación se establecieron con sesenta registros y semilla 20261026, mediante el generador predeterminado de NumPy: las horas se obtuvieron de una distribución uniforme entre cero y cuarenta y se redondearon a una decimal; después se generó un componente normal de media cero y desviación nueve, utilizado en la expresión puntuación = 45 + 1,1 × horas + componente aleatorio, con límites de cero y cien y un segundo redondeo a una decimal; el orden de esas operaciones forma parte del procedimiento; generar primero una puntuación y redondear las horas después produciría otro conjunto, aunque el texto mencionara la misma semilla y las mismas distribuciones, por lo que la reproducción necesita conservar la secuencia y las versiones efectivamente utilizadas, en este caso Python 3.12.14, NumPy 2.3.5, pandas 2.2.3 y SciPy 1.17.0.

El coeficiente 1,1 introduce deliberadamente una dependencia positiva en el generador; encontrar después una asociación del mismo signo permite comprobar cómo funciona el análisis sobre esa construcción, pero no aporta una estimación de cuánto cambia el aprendizaje cuando una persona estudia una hora adicional, porque las cifras provienen de una regla matemática y no de observaciones de estudiantes.

La tabla 7 presenta los resultados de la base depurada; se utilizaron los valores completos de los cálculos para obtener las medidas y se redondearon al comunicarlas, evitando que la precisión visual de la tabla se confunda con la precisión de una medición empírica que aquí no existe.

Tabla 7.

Descripción y asociación calculadas sobre los sesenta registros sintéticos

Medida	Horas por semana	Puntuación o asociación
Registros completos	60	60
Media	21,99 h	68,26 puntos
Desviación estándar muestral	10,33 h	14,58 puntos
Mediana	21,55 h	67,55 puntos
Mínimo–máximo	0,8–39,8 h	33,3–93,7 puntos
Pearson r	Entre ambas variables	0,8492
Spearman ρ	Entre ambas variables	0,8438

Recta ajustada	Intercepto 41,9000	Pendiente 1,1985 puntos/hora
R^2 dentro del archivo	Ajuste descriptivo	0,7212

Las horas tienen una media de 21,99 y una mediana de 21,55, mientras la puntuación alcanza 68,26 y 67,55, respectivamente; esa cercanía ayuda a describir el centro de este archivo, aunque por sí sola no demuestra normalidad ni descarta observaciones influyentes, y debe leerse junto con la dispersión y el gráfico de los pares, donde puede examinarse cómo se comportan las puntuaciones a lo largo del rango de dedicación simulado.

Para la asociación se calculó Pearson, que resume la relación lineal, y se utilizó Spearman como contraste descriptivo basado en rangos; los valores fueron 0,8492 y 0,8438, de modo que ambas medidas conservaron signo y magnitud próximos en este conjunto, coincidencia que informa sobre las representaciones examinadas y no prueba que la relación permanezca igual bajo otros generadores, otras escalas o datos educativos reales; mantener las dos medidas permite inspeccionar una decisión analítica previamente declarada, pero probar muchas alternativas y comunicar solo la que resulte más alta produciría otra clase de selección; el registro debe mostrar qué contraste se planificó y qué análisis surgió después de observar los valores.

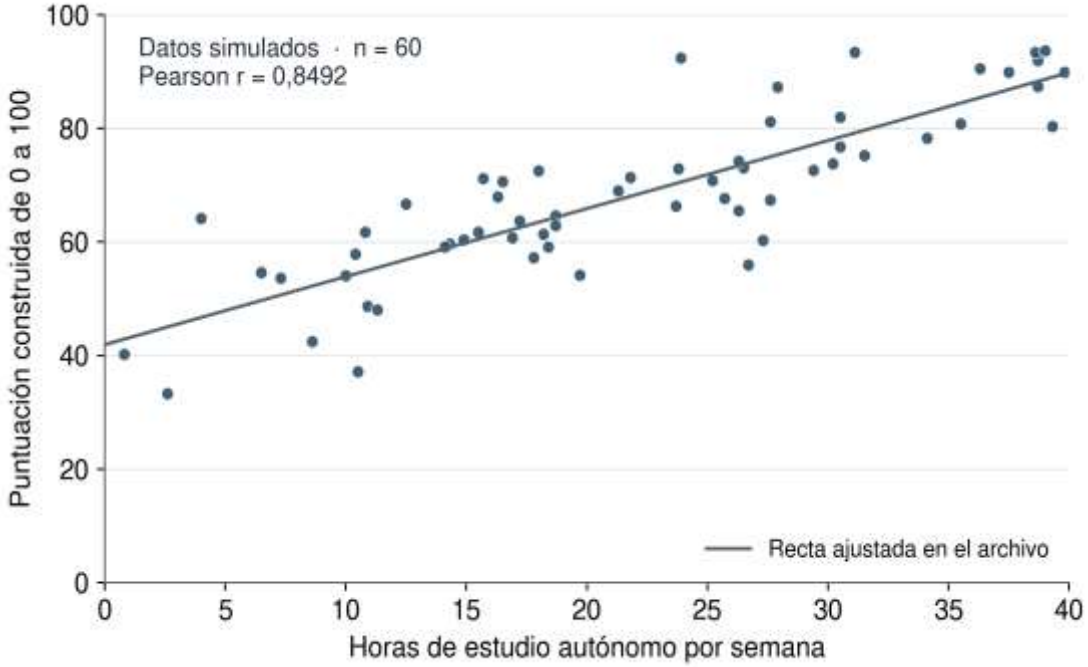
La documentación de SciPy (2026) especifica que pearsonr devuelve el coeficiente y una prueba de ausencia de correlación bajo determinados supuestos, distinción relevante porque una función puede producir más resultados de los que corresponde interpretar; aquí se comunica el coeficiente como descripción del archivo y se omite la lectura inferencial del valor p, ya que el ejemplo no contiene una muestra empírica sobre la cual sostener una conclusión poblacional.

La recta ajustada fue puntuación estimada = 41,9000 + 1,1985 × horas, con R^2 = 0,7212 dentro del archivo; su pendiente expresa cómo se ordena el ajuste entre los pares disponibles y su proporción de variación explicada se refiere exclusivamente a esa representación matemática, por lo que sería incorrecto convertirla en un porcentaje del rendimiento causado por estudiar o presentarla como rendimiento predictivo ya comprobado fuera de la base utilizada.

La figura 4 permite inspeccionar los pares y la recta ajustada; los puntos individuales y los límites de las escalas conservan información que una línea aislada ocultaría.

Figura 4.

Horas de estudio autónomo y puntuación en el archivo sintético



El artículo de Waskom (2021) presenta una biblioteca de visualización estadística orientada a relacionar estructura de los datos y representación gráfica, recordando que graficar forma parte del análisis; en este capítulo se utiliza una representación de dispersión producida con Matplotlib, bajo la misma paleta del libro, porque el criterio decisivo es que las marcas correspondan a los valores ejecutados y que ejes, unidades y escala permitan examinarlos, con independencia de la biblioteca elegida.

La comprobación del código debe recorrer las entradas y las operaciones antes de revisar el resultado final; una instrucción que selecciona todas las columnas numéricas podría incluir el identificador convertido en entero, otra podría redondear demasiado pronto y una tercera omitir registros incompletos sin comunicar cuántos desaparecen, errores que una salida numéricamente plausible no permite distinguir sin examinar qué filas y qué variables llegaron al

cálculo; el prompt C3-P1 delimita la tarea que podría encomendarse a un asistente, aunque la corrección de la ejecución se determina mediante el código y sus controles; se conserva como instrucción reutilizable y no como transcripción de una prueba comparativa entre modelos.

Prompt C3-P1 · Proponer un análisis comprobable

> Trabaja con 60 registros sintéticos, identificadores S001–S060, horas_semana de estudio autónomo entre 0 y 40 y puntuacion entre 0 y 100. Propón código Python para comprobar identificadores, unidades, rangos y ausencias antes de describir ambas variables y calcular Pearson y Spearman. Separa descripción, asociación y predicción. Conserva un registro de cualquier cambio; no sustituyas ausencias ni elimines extremos sin una regla justificada. Añade una segunda implementación de Pearson que relea el CSV y utilice la fórmula con sumas. No inventes cifras ni declares una ejecución si no dispones de su salida.

Estado y procedencia. Instrucción didáctica propuesta el 26 de octubre de 2026 para reutilización; no se presenta como un intercambio externo ejecutado ni como una evaluación de un modelo identificado, y el código que sigue sí fue ejecutado en el entorno declarado.

El siguiente código fue ejecutado durante la elaboración del capítulo y reproduce la base limpia y las operaciones principales, incluyendo una lectura independiente del archivo exportado mediante la biblioteca estándar; sus resultados son los informados en la tabla 7, mientras las incidencias de preparación se documentan por separado para mantener visible la diferencia entre generar el escenario y corregir su copia de trabajo.

```
from pathlib import Path
import csv
from decimal import Decimal, getcontext
import numpy as np
import pandas as pd
from scipy import stats

# Base sintética; orden de generación fijado.
rng = np.random.default_rng(20261026)
n = 60
x = np.round(rng.uniform(0, 40, n), 1)
error = rng.normal(0, 9, n)
y = np.round(np.clip(45 + 1.1*x + error, 0, 100), 1)
base = pd.DataFrame({
    "id": [f"S{i:03d}" for i in range(1, n+1)],
    "horas_semana": x, "puntuacion": y})
assert base.id.is_unique and not base.isna().any().any()
assert base.horas_semana.between(0, 40).all()
assert base.puntuacion.between(0, 100).all()
base.to_csv("base_analitica.csv", index=False,
            float_format="%.1f")
r = stats.pearsonr(x, y).statistic
rho = stats.spearmanr(x, y).statistic
ajuste = stats.linregress(x, y)
```

```
print("r, rho:", r, rho)
print("intercepto, pendiente:",
      ajuste.intercept, ajuste.slope)

# Segunda vía: CSV + Decimal + fórmula explícita.
getcontext().prec = 40
with open("base_analitica.csv", newline="") as f:
    filas = list(csv.DictReader(f))
xd = [Decimal(a["horas_semana"]) for a in filas]
yd = [Decimal(a["puntuacion"]) for a in filas]
nd = Decimal(len(xd))
sx, sy = sum(xd), sum(yd)
sxx = sum(a*a for a in xd)
syy = sum(a*a for a in yd)
sxy = sum(a*b for a, b in zip(xd, yd))
r_control = (nd*sxy-sx*sy) / (
    (nd*sxx-sx*sx)*(nd*syy-sy*sy)).sqrt()
assert abs(float(r_control)-r) < 1e-12
print("r por segunda vía:", r_control)
sensibilidad = [stats.pearsonr(
    np.delete(x, i), np.delete(y, i)).statistic
    for i in range(n)]
print("r al retirar uno:",
      min(sensibilidad), max(sensibilidad))
```

Salida comprobada: r = 0,8492297243905182; ρ = 0,8437843020919813; intercepto = 41,899992612497385; pendiente = 1,1984996538200376; sensibilidad entre 0,8383534355129604 y 0,8648973987555325.

La segunda vía de comprobación leyó los números decimales del CSV y aplicó la fórmula de Pearson con Decimal, sin utilizar la función estadística que produjo el primer coeficiente; la coincidencia fue inferior a una diferencia de 10^{-12}, con 0,8492297243905182 en la primera vía y 0,8492297243905181 al expresar la segunda en dieciséis decimales, resultado que respalda la consistencia aritmética y de transferencia del archivo, aunque ambos procedimientos conservan el mismo objeto de análisis y la misma definición de correlación; para hacer inspeccionable ese contraste se conservaron las sumas $\Sigma x = 1319{,}4$; $\Sigma y = 4095{,}3$; $\Sigma x^2 = 35308{,}02$; $\Sigma y^2 = 292061{,}33$ y $\Sigma xy = 97599{,}50$, con $n = 60$; esos componentes permiten reconstruir el numerador y el denominador sin depender de una captura de pantalla o de una afirmación del asistente sobre la exactitud de su propia respuesta.

La repetición de una función con los mismos argumentos sirve para reconocer estabilidad computacional, pero una comprobación independiente necesita introducir una vía que pueda revelar otro tipo de error; releer el archivo exportado permite advertir problemas de serialización o separación decimal, y aplicar una fórmula diferente en su implementación ayuda a identificar una llamada

incorrecta, aunque ninguna de esas operaciones comprueba si las horas representan adecuadamente el concepto de hábitos de estudio; en una auditoría más amplia también sería conveniente que otra persona revisara la selección de variables y las reglas de preparación; esa revisión no se ha simulado aquí, por lo que la independencia declarada corresponde a dos implementaciones computacionales y no a la intervención de dos equipos de investigación.

La evaluación predictiva exige controles adicionales; Kapoor y Narayanan (2023) documentan la fuga de información como un problema de reproducibilidad en ciencia basada en aprendizaje automático, mientras Lones (2024) desarrolla orientaciones para separar ajuste, selección y evaluación, aportes que convergen en impedir que el resultado que se pretende evaluar participe indebidamente en las decisiones del modelo; este problema puede aparecer al seleccionar variables sobre toda la base antes de dividirla o al utilizar reiteradamente la prueba para escoger la configuración más favorable.

En la documentación de scikit-learn (2026) se muestran transformaciones encadenadas con estimadores para que el ajuste se realice sobre entrenamiento y la prueba reciba únicamente la transformación ya aprendida; esa organización reduce errores operativos frecuentes, aunque la separación debe respetar también la unidad del estudio, porque varias observaciones de la misma persona repartidas entre grupos pueden conservar dependencias que una división aleatoria por filas no reconoce.

El caso no entrena un predictor y, por esa razón, no presenta una partición de entrenamiento y prueba como un trámite obligatorio para calcular su asociación descriptiva; si se ampliara hacia predicción, habría que formular un objetivo nuevo, decidir qué variables estarán disponibles al anticipar la puntuación y establecer cómo se evaluará el desempeño en unidades no utilizadas para ajustar ni seleccionar el modelo.

El examen de Varoquaux y Cheplygina (2022) sobre aprendizaje automático en imágenes médicas identifica problemas de evaluación y recomienda comparaciones y condiciones de prueba rigurosas, argumento que debe conservar su procedencia clínica; su utilidad metodológica para el ejemplo educativo reside en exigir un referente pertinente y una evaluación coherente

con el uso previsto, sin trasladar resultados de imágenes médicas como si midieran la eficacia de una herramienta para analizar estudiantes.

3.2.2. Análisis cualitativo y revisión de categorías e interpretaciones

En un análisis cualitativo, las palabras conservan relaciones con quien habla, con la pregunta formulada y con el contexto del intercambio; solicitar una lista de temas puede ayudar a explorar el material, pero la revisión necesita recuperar los pasajes que sostienen cada propuesta y distinguir una expresión literal de una inferencia, porque una etiqueta general puede sonar apropiada mientras elimina una condición que modifica el sentido del relato.

El ejemplo trabajado por Byrne (2022) sitúa el análisis temático reflexivo dentro de decisiones epistemológicas y del trabajo interpretativo del investigador, mostrando por qué los enfoques de análisis temático no deben tratarse como equivalentes; desde esa perspectiva, exigir siempre coincidencia entre codificadores como única medida de calidad puede ser incongruente con el método declarado, aunque otros diseños que utilizan un libro de códigos sí puedan justificar controles de consistencia definidos de antemano.

Esta diferencia afecta la tarea que se asigna a la IA; aplicar categorías establecidas requiere examinar la correspondencia de cada fragmento con las reglas de inclusión, mientras explorar patrones exige mantener abierta la posibilidad de reformular categorías y reconocer relaciones no previstas, de modo que el mismo prompt no debería anunciar simultáneamente una codificación cerrada y una interpretación enteramente inductiva sin explicar cómo se resolverá esa tensión; un libro de códigos necesita definiciones y ejemplos que permitan reconocer cuándo una etiqueta corresponde y cuándo debe evitarse; sin esa información, pedir consistencia puede reducirse a repetir nombres de categorías, mientras el criterio de asignación cambia silenciosamente entre fragmentos.

En su exploración de dos conjuntos previamente analizados, Morgan (2023) encontró mayor facilidad para recuperar contenidos descriptivos que temas interpretativos más sutiles, señalando además la limitación de comparar las

respuestas con análisis que él mismo conocía; el resultado respalda examinar qué nivel de elaboración aporta la herramienta y evita atribuirle una capacidad interpretativa general a partir de coincidencias que dependen de la tarea, las preguntas y el conocimiento previo de quien evalúa.

Para enseñar esa revisión se construyeron cinco fragmentos ficticios relacionados con el estudio autónomo; se identifican como Q01 a Q05 y no se enlazan con S001 a S060, porque no proceden de entrevistas a las unidades del archivo numérico ni permiten explicar sus puntuaciones, de manera que su proximidad temática sirve para comparar procedimientos analíticos sin simular un diseño mixto que nunca se ejecutó.

> Q01. «Puedo sentarme dos horas, pero me interrumpen cuando cuido a mi hermano; a veces termino sin saber cuánto estudié de verdad.»
>
> Q02. «Antes releía todo y no entendía; ahora hago ejercicios y comparo dónde me equivoco, aunque algunos días estudio menos.»
>
> Q03. «Cuando no hay señal descargo los materiales en casa de un amigo y los reviso después; lo difícil es preguntar cuando no entiendo.»
>
> Q04. «Le dediqué casi toda la tarde a una unidad y seguí confundido; el tiempo estuvo, pero necesitaba otra explicación.»
>
> Q05. «El teléfono me sirve para buscar una fórmula; si dejo las notificaciones abiertas, me distraigo y pierdo el hilo.»

Q01 describe una interrupción y sus efectos sobre la continuidad del estudio, mientras Q02 propone una diferencia entre cantidad de tiempo y estrategia; Q03 añade un condicionante de acceso y una acción para afrontarlo, Q04 introduce dificultad pese a una dedicación prolongada y Q05 limita la supuesta utilidad del teléfono a una tarea concreta, variedad que impide agrupar todas las expresiones bajo una valoración uniforme del compromiso académico sin perder contenido relevante.

La tabla 8 presenta una codificación inicial elaborada para el ejercicio; cada fila enlaza una etiqueta con un segmento identificable y con un límite interpretativo, de modo que el lector pueda examinar la decisión sin asumir que los cinco fragmentos constituyen evidencia empírica o que las categorías han sido

validadas por participantes o expertos; los límites de la última columna son parte de la decisión: impiden convertir una dificultad de acceso en un diagnóstico de exclusión permanente o una estrategia puntual en un rasgo estable de personalidad, inferencias que requerirían un material y una justificación adicionales.

Tabla 8.

Codificación inicial y límites de interpretación de los fragmentos ficticios

ID	Código propuesto	Anclaje en el fragmento	Límite de la interpretación
Q01	Continuidad interrumpida	Interrupciones durante el cuidado del hermano	No informa motivación ni calificación
Q02	Cambio de estrategia	Resolver ejercicios frente a releer	Autovaloración ficticia; sin medición de eficacia
Q03	Acceso condicionado y adaptación	Descarga previa por falta de señal	No demuestra desconexión permanente
Q04	Dificultad persistente	Horas dedicadas sin comprender una unidad	No permite diagnosticar capacidad ni estimar puntuación
Q05	Uso ambivalente del teléfono	Consulta puntual e interrupción posterior	No prueba un efecto causal del dispositivo

En De Paoli (2024), la exploración con dos conjuntos abiertos de entrevistas muestra posibilidades de recuperar temas mediante un modelo de lenguaje y mantiene un carácter exploratorio sobre los límites del enfoque; su contribución permite discutir el procedimiento y sus condiciones, pero no convierte la similitud con análisis anteriores en garantía de que toda formulación temática producida por IA haya identificado la interpretación mejor respaldada por el contexto.

El prompt C3-P2 pide conservar la separación entre texto, código e inferencia; también permite que un fragmento reciba más de una etiqueta y prohíbe atribuir diagnósticos o rendimientos no informados, condiciones que hacen revisable la tarea sin prometer que el sistema cumplirá automáticamente cada restricción.

Prompt C3-P2 · Revisar códigos sin agregar información

Analiza únicamente los cinco fragmentos ficticios Q01–Q05 reproducidos en este apartado. Propón códigos iniciales vinculados con condiciones y estrategias del estudio autónomo; admite más de un código cuando el texto lo justifique. Para cada código, conserva el identificador y un fragmento literal que lo respalde. Distingue descripción explícita e inferencia; señala ambivalencias o pasajes que cuestionen el

patrón. No atribuyas diagnósticos, calificaciones ni efectos causales, no declares saturación y no relaciones estos textos con los registros S001–S060. Devuelve una propuesta revisable, no una interpretación validada.

Estado y salida mostrada. Prompt propuesto el 26 de octubre de 2026; la tabla 8 es una elaboración editorial ilustrativa y la respuesta defectuosa del párrafo siguiente es ficticia, sin ejecución externa ni comparación de versiones de modelos.

Una respuesta ficticia defectuosa podría afirmar que «los cinco estudiantes presentan baja motivación y el teléfono causa su bajo rendimiento»; el error no se limita a elegir nombres de categorías demasiado amplios, pues añade una característica ausente, inventa un resultado académico y transforma una mención funcional del dispositivo en un mecanismo causal, por lo que corregir solo la redacción dejaría intacta la falla de evidencia.

La revisión editorial de esa muestra conserva el reconocimiento de interrupciones en Q01 y de conectividad en Q03, retira el diagnóstico sobre motivación y mantiene la ambivalencia de Q05; la decisión se basa en los fragmentos completos y no en el número de veces que la IA repite una interpretación, dejando constancia de qué propuesta se modificó y qué material permite sostener la formulación aceptada.

El estudio de Mathis et al. (2024) comparó temas generados con un modelo abierto ejecutado localmente y temas producidos mediante análisis humano en entrevistas clínicas, obteniendo semejanzas moderadas a sustanciales según sus evaluaciones; ese resultado corresponde a un entorno y a procedimientos específicos, por lo que aporta un antecedente sobre comparación de productos y no una certificación de equivalencia entre un generador y un investigador en cualquier corpus, idioma o marco interpretativo.

Disponer de varios temas parecidos tampoco informa por sí solo si se han conservado los casos que contradicen el patrón; aquí Q04 impide reducir el ejercicio a «más tiempo, mejor resultado», porque expresa una dificultad que persiste pese a estudiar durante varias horas, aunque no indique una puntuación con la que pudiera contrastarse estadísticamente, razón por la cual debe preservarse como matiz interpretativo y no convertirse en una supuesta excepción cuantificada.

La frecuencia de una expresión puede orientar la inspección, pero una experiencia mencionada una sola vez puede ser relevante para la pregunta; tampoco corresponde utilizar estos cinco fragmentos para declarar saturación, puesto que fueron escritos para una demostración y no emergieron de un proceso de muestreo y análisis, de modo que ampliar la cantidad mediante generación automática aumentaría el material didáctico disponible sin aportar diversidad de experiencias observadas; un fragmento infrecuente también puede cuestionar la forma en que se formuló una categoría general; su relevancia dependerá de lo que permita comprender sobre el problema, por lo que contar apariciones y argumentar su significado deben conservarse como operaciones relacionadas que responden a preguntas diferentes.

La perspectiva de Lee et al. (2024) examina aplicaciones de ChatGPT en codificación, construcción de temas y preparación de citas, identificando pérdidas de detalle y modificaciones textuales que requieren revisión; su aporte es especialmente pertinente para distinguir una síntesis del analista de las palabras atribuidas a un participante, porque corregir un fragmento para que resulte más fluido puede alterar precisamente la condición o la duda que el análisis debía conservar.

Una matriz cualitativa útil permite volver al texto y reconocer cambios de interpretación; guardar únicamente el nombre final de la categoría vuelve invisible el trabajo que la produjo, mientras conservar identificador, fragmento, código inicial, revisión y razón del cambio facilita discutir discrepancias con el equipo, especialmente cuando una propuesta automatizada desplaza un concepto del marco teórico o reúne experiencias distintas bajo una denominación que parece neutral; cuando dos interpretaciones compiten, recuperar el mismo pasaje permite discutir qué parte es explícita y qué parte depende del marco de lectura; la discrepancia puede motivar una reformulación más precisa sin obligar a escoger automáticamente la categoría más frecuente o la propuesta con redacción más convincente.

La comprobación cualitativa termina en una argumentación que el material pueda sostener; para este ejercicio resulta defendible afirmar que los textos construidos representan distintas condiciones y estrategias vinculadas con el

estudio, pero no estimar cuántos estudiantes reales viven esas situaciones ni utilizar sus expresiones para explicar la correlación de la base numérica, cuyos valores se generaron mediante un procedimiento independiente.

3.3. Contraste de resultados y explicación de errores en el caso práctico

El contraste reúne los productos y examina si hablan del mismo archivo, del mismo procedimiento y del mismo alcance; la entrada con incidencias ofrece una advertencia concreta, porque un cálculo automático que descartara la puntuación vacía y aceptara lo demás devolvería sesenta filas utilizables, cifra idéntica al tamaño previsto, aunque esas filas incluirían un duplicado y representarían solo 59 identificadores diferentes.

En esa ruta incorrecta, la correlación calculada fue 0,1511, frente a 0,8492 después de las reparaciones documentadas; la discrepancia resulta de combinar la unidad inconsistente, la puntuación fuera de rango, la ausencia y la repetición, por lo que esta comparación no permite atribuir todo el cambio a una sola incidencia, pero sí demuestra que conservar el número esperado de filas no garantiza estar analizando las unidades y las medidas previstas.

Tampoco correspondería describir la corrección como una mejora de la relación estadística; las reglas se justifican por el diccionario y por la procedencia del dato, y habrían debido aplicarse aunque el coeficiente disminuyera, porque usar la fuerza de la asociación para decidir qué transformación aceptar invertiría el orden de la evaluación y permitiría construir un resultado favorable mediante ajustes posteriores.

La tabla 9 confronta productos y comprobaciones sin atribuir a un modelo respuestas que no se obtuvieron; diferencia la ejecución de una ruta informática defectuosa, la salida calculada sobre la base corregida y las afirmaciones inventadas para el ejercicio, distinción necesaria para aprender de un error reproducible sin convertir un ejemplo editorial en un registro experimental de desempeño de IA.

Tabla 9.

Contraste entre productos del ejercicio y procedimientos de comprobación

Producto y estado	Resultado o afirmación	Control efectuado	Decisión
Ruta defectuosa ejecutada	r = 0,1511; 60 filas tras descartar una ausencia	Unidad, rango e identificadores: solo 59 códigos distintos	Rechazar como análisis del protocolo
Base depurada ejecutada	r = 0,8492; 60 pares y 60 códigos	Diccionario, origen de correcciones y fórmula Decimal	Conservar como descripción del caso
Afirmación numérica ficticia	“Cada hora causa una mejora de 1,1985 puntos”	Examinar diseño y origen del coeficiente	Retirar lenguaje causal
Respuesta cualitativa ficticia	“Los cinco presentan baja motivación”	Volver a Q01–Q05 y buscar respaldo textual	Retirar diagnóstico ausente
Cruce interpretativo ficticio	“La conectividad explica la correlación”	Los textos Q y los datos S no tienen enlace empírico	Mantener separados los dos ejercicios

La estabilidad del coeficiente se examinó retirando una observación cada vez y recalculando Pearson sobre los 59 pares restantes; las sesenta ejecuciones produjeron valores entre 0,8384 y 0,8649, de manera que ninguna eliminación individual cambió el signo ni desplazó ampliamente la asociación dentro de este escenario, aunque el recorrido no representa un intervalo de confianza ni permite afirmar robustez frente a cambios de diseño, selección de población o generación de datos; se trata de una pregunta de sensibilidad acotada: cuánto depende el coeficiente observado de cada registro individual; responderla no autoriza a retirar la observación que produzca el valor más favorable, ni demuestra que sesenta registros sean suficientes para otro objetivo, y conserva su utilidad precisamente porque declara la perturbación examinada y comunica el conjunto del rango obtenido en lugar de escoger una ejecución para presentarla como definitiva.

Para reconocer la versión analizada se calculó además una huella SHA-256 del CSV depurado, cuyo prefijo es e6e8b0a7d046bcf; una huella idéntica informa que los bytes comparados coinciden, mientras una diferencia puede deberse incluso al formato de exportación, por lo que este control sirve para identificar

archivos y no acredita la calidad de las variables, la pertinencia del análisis o la autenticidad de los supuestos que hicieron posible la simulación.

El registro conserva los archivos de entrada y salida, la semilla, las versiones, las transformaciones y los resultados; cuando una revisión posterior cambie una regla, deberá producir una versión reconocible y actualizar los productos dependientes, porque modificar una puntuación en la base sin regenerar la tabla y la figura deja un manuscrito con resultados incompatibles, aunque cada componente pueda parecer correcto cuando se inspecciona por separado; también debe precisarse qué versión se utilizó para cada producto: un nombre como «resultados finales» no permite distinguir correcciones sucesivas, mientras una identificación estable vinculada con el archivo y con la ejecución facilita reconocer si la revisión obliga a sustituir una sola tabla o todo el análisis.

La revisión de una explicación también necesita controles propios; escribir que el tiempo de estudio «determina» el rendimiento introduce un mecanismo que el diseño no evaluó, mientras afirmar que la asociación «es explicada por dificultades de conectividad» utilizaría los fragmentos inventados para justificar números generados independientemente, dos errores distintos que no desaparecen al conservar intactos los coeficientes o al añadir una referencia general sobre tecnología educativa; la formulación aceptable del caso es más precisa: en los sesenta registros sintéticos, las horas y la puntuación mostraron una asociación positiva según las dos medidas examinadas; la dirección fue introducida en la regla de generación y el cálculo de Pearson se reprodujo mediante una fórmula independiente, por lo que el resultado enseña a ejecutar y comprobar un análisis, sin describir el comportamiento de una población universitaria ni estimar el efecto de una intervención educativa.

En la demostración cualitativa, la aceptación depende de que cada interpretación conserve vínculo con los fragmentos y reconozca las diferencias entre ellos; allí no existe un valor aritmético final que resuelva todas las discrepancias, pero sí puede examinarse si se atribuyó una frase inexistente, se omitió una condición relevante o se presentó como observación una inferencia del analista, controles que permiten discutir el producto con razones visibles.

Capítulo IV: Ética e integridad a lo largo de la investigación



Una base puede estar depurada y un resultado puede reproducirse sin que la investigación haya resuelto todavía sus problemas éticos: conocer quién autorizó el uso de la información, qué personas quedan representadas, quién puede impugnar una interpretación y qué ocurrió entre la primera instrucción y el resultado publicado exige examinar decisiones que no aparecen en una tabla de coeficientes, aunque condicionan su legitimidad y el alcance de lo que se comunica; el caso del capítulo anterior ofrece un terreno concreto para esta revisión porque su procedencia está delimitada: sesenta registros numéricos generados mediante un procedimiento explícito y cinco fragmentos cualitativos inventados para enseñar codificación; no hubo reclutamiento, entrevistas ni acceso a expedientes estudiantiles, de modo que el examen de este capítulo conserva esa condición y evita convertir una simulación documentada en un estudio de personas por efecto de la redacción. La pregunta cambia de lugar: además de comprobar qué hizo el procedimiento, hay que decidir qué estaba autorizado a hacer y qué afirmaciones permite sostener.

4.1. Responsabilidad del investigador y límites de la delegación

Delegar una tarea supone establecer su objeto, sus límites y el control que permitirá aceptarla; pedir a un asistente que ordene una matriz bibliográfica puede ser compatible con una revisión directa de cada referencia, mientras que encargarle la decisión de excluir participantes modifica el diseño y exige una justificación que debe quedar en manos de quienes responden por el estudio, incluso si ambos encargos caben en una instrucción de pocas líneas y producen una respuesta con apariencia semejante.

La responsabilidad científica tampoco se transfiere mediante una declaración genérica de uso de IA, pues reconocer que una herramienta intervino no explica qué datos recibió ni demuestra que sus propuestas fueran examinadas; Resnik y Hosseini (2025) sitúan la discusión en las obligaciones de los investigadores respecto de la integridad de los resultados, la procedencia de materiales y la transparencia sobre el uso de estas herramientas, un planteamiento que permite distinguir la colaboración técnica de la capacidad de responder por errores o daños.

En una investigación pequeña, esa capacidad puede concentrarse en una sola persona; aun así, conviene separar las funciones de solicitar una salida, contrastarla con evidencia y autorizar su incorporación, porque escribir tres veces el nombre del investigador en un formulario no crea independencia de revisión, pero sí puede obligarlo a hacer visibles los criterios con los que actuó y a identificar qué controles adicionales necesita antes de cerrar una decisión importante.

La supervisión tiene contenido cuando puede cambiar el resultado. Si el control se reduce a leer una respuesta y encontrarla convincente, la revisión sigue dependiendo de la misma fluidez que debía poner a prueba. Ante una propuesta de imputar un dato faltante, por ejemplo, el investigador necesita comprobar el origen de la ausencia y la regla del protocolo antes de aceptar un valor, mientras que una sugerencia de mejorar la sintaxis puede resolverse mediante una lectura que preserve el significado, sin imponer a ambas intervenciones un procedimiento idéntico ni atribuirles el mismo riesgo.

El trabajo de Hosseini et al. (2024) sobre escritura académica defiende la comprobación por especialistas del dominio y la responsabilidad de los autores sobre el texto que incorporan, incluyendo la exactitud factual y bibliográfica; su aportación se utiliza aquí como argumento editorial sobre la revisión competente, sin presentar aquella discusión como evidencia experimental del rendimiento de los modelos disponibles en 2026 ni como una norma obligatoria para todas las revistas.

En el caso simulado, aceptar el coeficiente calculado requiere saber de dónde provienen sus sesenta pares y cómo se obtuvo la cifra; aceptar una explicación de ese coeficiente demanda, además, reconocer que la asociación fue incorporada al mecanismo generador, por lo que una salida que atribuyera la puntuación a la voluntad, al talento o al uso de una aplicación tendría que rechazarse aunque estuviera escrita con precisión gramatical y acompañada de una explicación extensa.

La dependencia tecnológica aparece también cuando el equipo conserva archivos que puede abrir, pero ya no comprende las operaciones que los produjeron: una plantilla cerrada, una función automática sin parámetros visibles

o un historial que desaparece al cambiar de cuenta pueden volver costosa la revisión posterior, de ahí que la elección de herramientas deba considerar la posibilidad de exportar insumos, leer el procedimiento y continuar el trabajo cuando el servicio deje de estar disponible.

Stahl (2021) propone examinar la ética de la IA desde un ecosistema de actores y capacidades, enfoque que amplía la atención más allá del comportamiento individual del usuario hacia las condiciones institucionales y técnicas de su trabajo; aplicado a una investigación, permite preguntar quién facilita almacenamiento autorizado, quién forma a los equipos y quién interviene ante un incidente, sin diluir por ello la responsabilidad de quien decide utilizar un dato o publicar una conclusión.

La evidencia regional merece una lectura igualmente cuidadosa: Mancheno-Oviedo et al. (2026) describen percepciones sobre IA y pensamiento crítico mediante un cuestionario aplicado a 182 estudiantes de bachillerato, en un diseño no experimental; ese alcance admite discutir cómo el alumnado percibe sus prácticas, pero no demuestra que una herramienta produzca una mejora del pensamiento crítico ni autoriza a trasladar el resultado a estudiantes universitarios o al escenario sintético de este libro; la diferencia importa al redactar: afirmar que un grupo percibe apoyo para organizar ideas conserva el objeto de una respuesta de encuesta; afirmar que la IA desarrolla una capacidad introduce un efecto que requeriría otra evidencia, y el paso entre ambas formulaciones puede ocurrir durante una paráfrasis aparentemente menor, por lo que revisar los verbos forma parte del control metodológico y no debe dejarse para una corrección de estilo al final del manuscrito.

La tabla 10 organiza decisiones que atraviesan el estudio y señala qué evidencia permitiría revisarlas; los responsables consignados corresponden a funciones que habría que asignar en un proyecto real, mientras que en el ejercicio del libro solo se han realizado los controles técnicos y editoriales que se identifican expresamente en el apartado 4.3.2.

Tabla 10.

Matriz de riesgos, responsabilidades y evidencia revisable

Etapa	Riesgo	Medida	Función responsable	Evidencia
Diseño	Delegar decisiones sin criterio	Fijar finalidad y límites antes de usar IA	Investigador principal	Protocolo y regla de aceptación
Preparación	Comunicar datos innecesarios	Revisar procedencia y minimizar la entrada	Custodia de datos	Diccionario y autorización pertinente
Análisis	Aceptar errores o exclusiones arbitrarias	Comprobar unidades, rangos y reglas	Responsable de análisis	Código, salida y registro de cambios
Interpretación	Atribuir causas o rasgos ausentes	Contrastar cada inferencia con el diseño	Autor y revisión metodológica	Pasaje, dato y decisión argumentada
Comunicación	Presentar propuestas como hechos	Identificar estados y alcance de resultados	Autor responsable	Manuscrito y declaración de uso
Archivo	Perder versiones o exponer material	Definir accesos, conservación y recuperación	Custodia institucional	Inventario y permisos de acceso

El registro de una decisión no vuelve aceptable cualquier elección: si la tarea excede la competencia del equipo, si el dato carece de una procedencia autorizada o si la salida puede afectar a una persona sin un mecanismo de revisión, dejar constancia del problema debe conducir a detener o modificar esa parte del trabajo, en lugar de usar la existencia de una bitácora como argumento suficiente para continuar.

4.2. Protección de las personas y los datos

Proteger información comienza antes de elegir el servicio que la procesará, porque una solicitud de ayuda puede revelar más de lo que la operación necesita: enviar una entrevista completa para mejorar una categoría incorpora referencias familiares, circunstancias de salud o conflictos institucionales que quizá no intervengan en la pregunta analítica, y eliminar nombres al final de la conversación no deshace la comunicación que ya ocurrió al comienzo; la protección exige conocer el recorrido del material: dónde se obtuvo, qué copia

se transformó, quién puede consultarla, qué sistema externo recibiría una parte y durante cuánto tiempo se conservarían los archivos; sin ese mapa, el equipo puede tratar como una sola operación actividades que tienen finalidades, actores y condiciones distintas, desde la captura de una respuesta hasta el archivo de un borrador con comentarios que todavía contienen información identificable.

La referencia jurídica de este capítulo se sitúa en Perú, donde el artículo 2, numeral 6, de la Constitución protege la intimidad personal y familiar frente al suministro de información por servicios informáticos (Congreso Constituyente Democrático, 1993); este punto de partida permite relacionar las decisiones técnicas con los derechos de las personas, manteniendo los principios generales de protección como orientación para lectores de otros países, quienes deberán examinar la normativa aplicable a su propio proyecto.

La propuesta de Floridi (2023), en el planteamiento de su libro sobre ética de la IA, relaciona los principios con decisiones de diseño y gobernanza; esta perspectiva resulta útil para situar la protección dentro de la organización del estudio, donde una finalidad acotada, permisos diferenciados y una vía de corrección pueden modificar el tratamiento desde su inicio, aunque no sustituyen la identificación de las obligaciones jurídicas aplicables al contexto.

4.2.1. Marco peruano del consentimiento y la información sensible

En Perú, la Ley N.º 29733, Ley de Protección de Datos Personales, vincula el tratamiento con principios de finalidad, proporcionalidad y seguridad; los numerales 13.5 y 13.6 establecen las condiciones del consentimiento y su forma escrita para datos sensibles, mientras que el artículo 14 contempla supuestos en los que no se requiere consentimiento (Congreso de la República del Perú, 2011); en la práctica investigativa, esa lectura exige identificar el supuesto aplicable antes de utilizar expedientes o entrevistas, porque una autorización institucional de acceso no resuelve por sí sola la legitimidad de cada tratamiento posterior y la finalidad académica no permite presumir una excepción general para comunicar información a un asistente externo.

El Reglamento aprobado mediante Decreto Supremo N.º 016-2024-JUS precisa las características del consentimiento y la información que debe recibir su titular, incluidos los destinatarios y las transferencias previstas; sus artículos 1 a 8 permiten revisar qué se ha autorizado, y el artículo 10 establece una revocación sin efectos retroactivos (Presidencia de la República del Perú, 2024); un formulario comprensible debe corresponder al tratamiento que realmente se realizará, incluida la intervención de proveedores externos cuando proceda.

En un proyecto educativo real, un texto de información podría explicar qué tarea apoyará la IA, qué categorías de datos intervendrán, qué alternativas existen y ante quién se puede ejercer un derecho; aquí se trata de componentes que habría que desarrollar y revisar con la institución, porque un párrafo redactado por un asistente no acredita que el procedimiento haya sido aprobado ni que una persona lo haya comprendido o aceptado. El caso de este libro no tiene formularios firmados: no hubo participantes a quienes solicitarlos.

Conviene mantener esa precisión en todo el expediente, incluso cuando se adapten ejemplos para enseñar protocolos con personas, porque escribir «se obtuvo consentimiento informado» para completar la estructura de un informe introduciría un hecho inexistente; la simulación puede contener un modelo de documento, siempre identificado como propuesta, pero su utilidad didáctica depende de que el lector pueda distinguirlo de una actuación efectivamente realizada.

La Ley N.º 31814 incorpora el enfoque de riesgos, el desarrollo ético y la privacidad entre los principios para promover el uso de la inteligencia artificial; su artículo 1 sitúa esa promoción dentro del respeto de los derechos humanos (Congreso de la República del Perú, 2023); para una investigación, este marco permite examinar la herramienta junto con sus consecuencias para las personas, sin convertir el propósito de innovación en una autorización para tratar cualquier información ni atribuir a la ley una aprobación de los resultados producidos por un sistema.

El Reglamento de esa ley, aprobado mediante Decreto Supremo N.º 115-2025-PCM, distingue usos indebidos, de riesgo alto y de riesgo aceptable, y su artículo 26 remite a la normativa de protección de datos personales (Presidencia del

Consejo de Ministros, 2025); la clasificación depende del uso concreto y de sus consecuencias, por lo que apoyar una redacción y decidir sobre el acceso de una persona a un servicio requieren exámenes distintos; las obligaciones específicas deben leerse según el papel de cada actor y los plazos de implementación gradual, sin suponer que toda consulta académica constituye un sistema de riesgo alto.

En el ejercicio del libro, los valores se generaron sin registros de personas y los fragmentos cualitativos fueron inventados; esa procedencia permite delimitar el material revisado, pero no autoriza a llamar «sintética» a una transcripción real modificada ni a trasladar la misma conclusión a una base construida desde expedientes, porque cualquier cambio de insumos obliga a revisar la posibilidad de identificar personas y las condiciones del tratamiento antes de continuar.

El envío a servicios externos exige revisar también el flujo transfronterizo: los artículos 18 a 21 del reglamento de protección de datos desarrollan el nivel adecuado de protección, las garantías y la comunicación del flujo a la autoridad (Presidencia de la República del Perú, 2024); para evaluar un proveedor, el equipo necesita conocer quién recibe la información, dónde se trata y qué condiciones rigen su uso, sin dar por resuelto ese examen mediante una casilla de «uso académico».

La revisión normativa se cierra al 28 de octubre de 2026 y toma estas disposiciones como referencia para el contexto peruano; un proyecto real deberá comprobar las normas complementarias y las condiciones institucionales que correspondan al momento de su ejecución, pues el examen del caso didáctico no acredita cumplimiento jurídico para un tratamiento diferente.

Para avanzar con prudencia operativa, el equipo puede trabajar primero con un diccionario sin registros individuales y solicitar una revisión de los campos; esa minimización debe ocurrir antes del envío, con una persona responsable de decidir qué sale del entorno autorizado, porque pedir al mismo asistente que elimine información de un archivo que ya recibió desplaza la protección a una fase posterior a la exposición que se quería evitar.

Prompt C4-P1 · Detectar información que no debe enviarse

> Voy a utilizar un asistente externo para apoyar una investigación educativa. A partir del diccionario que aporto, identifica qué variables podrían revelar identidad o información sensible y propone minimización, agregación o sustitución. No solicites nombres, números de identificación ni expedientes reales. Diferencia anonimización de seudonimización y señala qué requiere revisión institucional.

Estado y salida. Instrucción reutilizable conservada como v1 el 27 de octubre de 2026; la respuesta siguiente es una elaboración editorial aplicada al diccionario declarado, sin ensayo comparativo de modelos ni revisión institucional; no se incorporaron registros de personas reales.

La entrada utilizada para desarrollar la respuesta editorial se limita al diccionario del caso: id identifica una fila inventada entre S001 y S060, horas_semana contiene un valor construido de 0 a 40 y puntuacion un valor construido de 0 a 100; Q01–Q05 son textos ficticios sin correspondencia con esas filas, de modo que no se adjuntaron nombres, matrículas, expedientes ni un conjunto de personas al que pudiera enlazarse el material.

Respuesta editorial aplicada al diccionario: no se identifica información de personas reales en los insumos declarados; para discutir una futura réplica empírica bastaría inicialmente con definiciones, escalas y rangos, y cualquier incorporación de códigos vinculables a estudiantes, relatos biográficos o resultados individuales exigiría revisar finalidad, permisos y condiciones institucionales antes de compartirlos; la sustitución de un nombre por un código no acredita anonimización cuando existe una clave o una combinación de atributos que permite restablecer la identidad. La tabla 11 distingue esa entrada efectivamente utilizada de los campos que podrían aparecer en una investigación real; sus medidas son propuestas de minimización, no una certificación de anonimato ni una autorización para transferir registros.

Tabla 11.

Minimización de la entrada: caso utilizado y posible adaptación empírica

Material	Estado en este ejercicio	Riesgo si procediera de personas	Entrada mínima propuesta
S001–S060	Códigos de filas inventadas; sin clave personal	Vinculación si existiera una clave o cruce	Definición del identificador; omitir filas para revisar el esquema

Horas y puntuación	Valores generados, sin expediente de origen	Inferencia por combinaciones o grupos pequeños	Escalas y rangos; agregados solo tras revisar su riesgo
Q01–Q05	Fragmentos ficticios y separados de la base numérica	Reconocimiento por detalles del relato	Descripción de la tarea; extractos minimizados si están autorizados
Nombres o matrículas	No existen en los insumos	Identificación directa y comunicación innecesaria	No solicitarlos para revisar el diccionario
Información sensible	No se incorporó información real de esta clase	Afectación de intimidad y derechos	Excluirla de la consulta inicial y revisar necesidad y garantías

Los textos libres requieren una atención distinta a la de los campos cerrados: una carrera poco frecuente, una fecha concreta y la descripción de una situación familiar pueden formar una combinación reconocible aunque ningún elemento aislado parezca suficiente, por lo que el equipo debe valorar qué conservar para responder la pregunta y qué sustituir por una descripción menos específica, dejando trazabilidad de la transformación en un archivo de acceso restringido cuando trabaje con material real; la eliminación tampoco debe hacerse de manera indiscriminada, porque suprimir toda referencia contextual puede vaciar el significado analítico de una experiencia; la decisión justificable consiste en conservar lo necesario dentro del entorno autorizado, reducir lo que se comunica a terceros y explicar qué pérdidas produce esa reducción, con una regla que pueda revisarse cuando cambie el propósito del análisis o el público del producto final.

4.2.2. Sesgos algorítmicos y representación de las poblaciones

La privacidad responde una parte de la protección, pero deja abierta la pregunta por las personas a quienes una investigación describe mal o excluye de sus conclusiones. Un conjunto sin nombres puede representar solo a quienes tuvieron conexión estable, respondieron en una lengua dominante o completaron una actividad en condiciones particulares, y la ausencia de identificadores directos no corrige ese desajuste entre la población declarada y la información que realmente se reunió.

La revisión de Mehrabi et al. (2021) ordena distintas fuentes de sesgo y definiciones de equidad en aprendizaje automático, mostrando que el problema atraviesa la construcción de datos y la evaluación de resultados; su taxonomía ayuda a formular preguntas de revisión, pero no proporciona una medida única que permita declarar «justa» cualquier aplicación sin especificar qué resultado se compara, entre qué grupos y con qué consecuencias para el uso previsto.

En el caso sintético no se dispone de grupos empíricos ni de una población muestreada: las sesenta filas permiten verificar operaciones y observar las consecuencias de errores introducidos, pero no estimar diferencias por género, territorio, discapacidad o condición socioeconómica; añadir esas etiquetas mediante una asignación aleatoria solo produciría otro ejercicio diseñado y tendría que identificarse como tal, sin presentarlo como una evaluación de desigualdades existentes. La ausencia de evidencia sobre discriminación no es una medición de equidad.

El argumento de Hooker (2021) cuestiona la idea de que el sesgo sea exclusivamente un problema de datos y señala la participación de las elecciones de diseño del modelo en los daños posibles; para el investigador, esta advertencia significa que aumentar el tamaño de una base no resuelve por sí mismo problemas asociados al objetivo optimizado, a los criterios de error aceptables o a la forma en que una salida se convierte en una decisión sobre personas.

Supóngase, como extensión hipotética, que una institución quisiera utilizar horas de estudio para priorizar apoyo: una regla que interprete menos horas como falta de compromiso podría perjudicar a quienes trabajan o cuidan a otras personas, mientras que otra que asigne recursos solo a las puntuaciones más bajas podría ignorar necesidades de acceso; revisar el criterio exige discutir la finalidad del apoyo y los errores tolerables antes de escoger un umbral o pedir a una IA que lo recomiende.

Birhane (2021) propone una ética relacional que atiende a contextos y comunidades afectadas, frente a una lectura de la justicia reducida a propiedades técnicas aisladas; esa contribución permite plantear la participación de las personas concernidas en la definición del problema, sin atribuir a una

consulta simbólica el poder de validar cualquier sistema ni suponer que un equipo pueda representar por anticipado todos los intereses del grupo estudiado; en un proyecto real, esa participación tendría momentos y consecuencias visibles: revisar la claridad de una pregunta, identificar categorías que resultan estigmatizantes o discutir qué información debería volver a la comunidad puede modificar el protocolo; si el equipo recoge observaciones y mantiene cada decisión intacta, necesitará explicar por qué, porque la mera celebración de una reunión no demuestra que las personas hayan influido en el diseño o que acepten el uso posterior del resultado.

Los sistemas de lenguaje introducen otra capa de representación: Bender et al. (2021) advierten, en su artículo de congreso, que el volumen de los corpus no garantiza diversidad de perspectivas y que las voces dominantes pueden reforzarse en esos materiales; esta discusión ayuda a revisar categorías, expresiones y ejemplos producidos por asistentes, sin asumir que una respuesta fluida en español refleja adecuadamente las variedades lingüísticas o los marcos culturales de una población peruana concreta.

La accesibilidad debe examinarse en el propio procedimiento de investigación, desde la posibilidad de responder sin un dispositivo reciente hasta la comprensión de las instrucciones y la disponibilidad de formatos compatibles con apoyos. Si el protocolo obliga a usar una cuenta de pago o una interfaz que parte de la población no puede manejar, el problema afecta tanto a las oportunidades de participación como a la composición de la evidencia que luego se analizará. También el equipo necesita una alternativa de trabajo cuando el servicio falla: conservar el diccionario en formato legible, poder ejecutar los cálculos sin depender de una conversación y mantener criterios de codificación comprensibles reduce la dependencia de una plataforma particular. Estas medidas no aseguran por sí mismas inclusión social, pero permiten identificar y corregir barreras que un relato centrado únicamente en la precisión del modelo dejaría fuera del examen.

4.3. Trazabilidad y revisión de las decisiones

La trazabilidad conecta una afirmación con el material que la sostiene y con las decisiones que hicieron posible incorporarla: si el lector llega al coeficiente

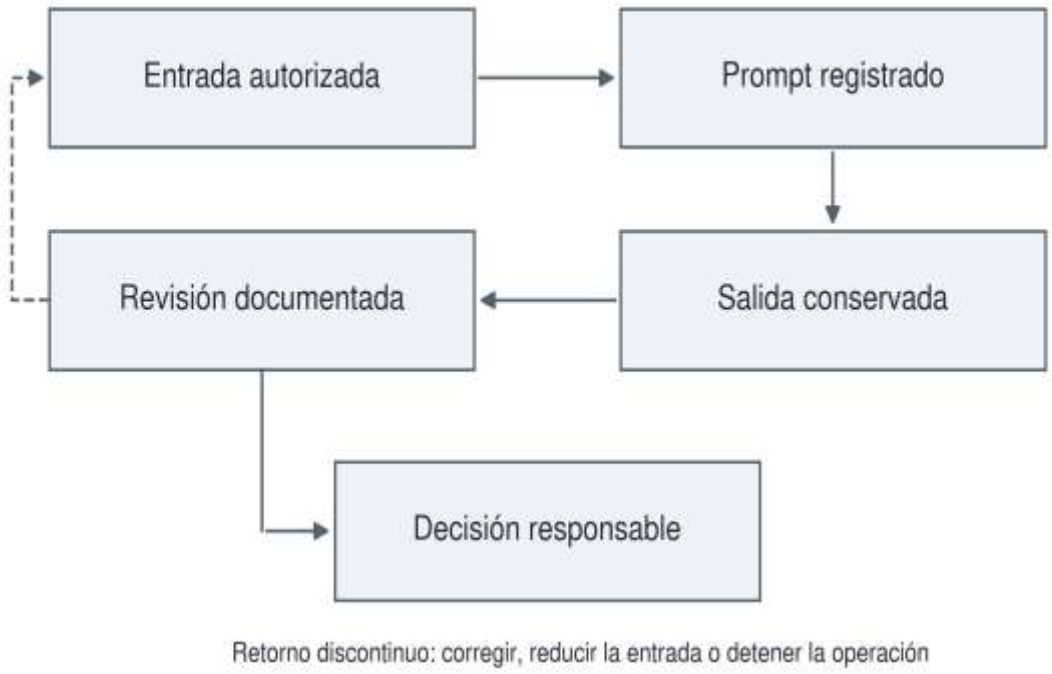
publicado, debería poder identificar el archivo analítico, la transformación de las unidades y la regla de exclusión aplicada, mientras que quien revise una interpretación cualitativa necesitará volver al fragmento y distinguir lo que estaba escrito de lo que propuso el analista o el asistente.

Conservar todo sin criterio tampoco facilita esa revisión, porque un historial de cientos de mensajes sin versiones ni estados puede ocultar la decisión importante entre intentos descartados; el expediente debe permitir reconstruir el recorrido pertinente y proteger al mismo tiempo los materiales restringidos, con un índice que separe insumos, instrucciones, salidas, controles y decisiones de publicación sin borrar las relaciones entre ellos.

La figura 5 sitúa esos elementos dentro de un ciclo que admite regresar a la entrada cuando la revisión detecta un problema; el retorno puede significar reducir información, reformular una instrucción o cancelar una operación, y no implica que toda salida deba aceptarse después de una edición superficial.

Figura 5.

Ciclo de trazabilidad de una tarea asistida por IA



4.3.1. Registro de herramientas, versiones, instrucciones y modificaciones

Un registro útil comienza con la identificación de la actividad: fecha, objetivo, persona o función responsable, herramienta y versión cuando se conocen, entrada autorizada, instrucción exacta, salida conservada y criterio de aceptación. Cuando algún dato no está disponible, como el identificador técnico de un modelo ofrecido mediante una interfaz, debe consignarse esa limitación en lugar de completar el campo con una versión supuesta que después impediría interpretar correctamente el expediente.

Hosseini et al. (2023) argumentan que la transparencia sobre herramientas de escritura debe permitir comprender su intervención en el trabajo académico y discuten formas de hacerla visible en el cuerpo del texto y en materiales complementarios; esta posición respalda una descripción de tareas y contribuciones, aunque la forma concreta de declarar el uso debe adecuarse a las instrucciones del destino editorial y no deducirse de una política universal inexistente.

Para este capítulo se registra el uso de un asistente de IA en la redacción y organización editorial, junto con búsquedas de fuentes originales y ejecución local de controles mediante código; no se dispone de un identificador técnico verificable del modelo conversacional ni de una configuración completa de muestreo, por lo que la documentación conserva las instrucciones publicadas y las salidas editoriales pertinentes, sin prometer que otra conversación reproducirá literalmente cada párrafo; las versiones deben indicar qué cambió y por qué: corregir una unidad, reducir una entrada para proteger información y retirar una afirmación causal son modificaciones de naturaleza distinta, de modo que guardar archivos llamados «final», «final2» y «definitivo» no basta para explicar el recorrido; un identificador estable, una fecha y una descripción breve del cambio permiten comparar estados sin tener que inferir la intención del autor a partir del nombre del documento.

En el ejercicio se distinguen la copia con incidencias deliberadas, la base analítica depurada del capítulo 3 y el registro de auditoría C4-A01 elaborado el 27 de octubre de 2026; el último añade controles y decisiones editoriales, pero

no constituye una nueva muestra ni modifica los sesenta pares, y el archivo de datos mantiene la misma huella digital que el utilizado para publicar los resultados previos.

La tabla 12 muestra cómo registrar los estados de las instrucciones y de los productos vinculados; las versiones de prompt designan instrucciones conservadas en el libro y no una comparación experimental entre modelos, mientras que el cambio de estado de los datos corresponde a operaciones que sí fueron ejecutadas y documentadas.

Tabla 12.

Versiones y estados documentados de los productos del caso

Registro	Entrada o producto	Cambio y motivo	Control o estado
C3-D0	Copia con incidencias deliberadas	Duplicado, unidad mixta, ausencia y valor inválido para enseñar control	Ejecutada; conservar como ruta defectuosa, no resultado aceptado
C3-D1	base_analitica.csv; 60 pares	Correcciones desde el origen sintético documentado	Ejecutada; id, rangos, ausencias y cálculo por dos vías
C3-P1 v1	Instrucción de análisis	Sin nueva versión de prompt en este capítulo	Propuesta publicada; el código vinculado sí se ejecutó
C3-P2 v1	Instrucción de codificación de Q01–Q05	Sin nueva versión de prompt en este capítulo	Propuesta; tabla de códigos editorial, sin validación externa
C4-P1 v1	Diccionario sin datos personales reales	Revisión de minimización añadida el 27-10-2026	Instrucción reutilizable y respuesta editorial aplicada
C4-P2 v1 / C4-A01	Registro de decisiones y CSV	Auditoría añadida; los datos permanecen iguales	Cálculo y huella ejecutados; evaluación ética editorial

El acceso a los registros puede tener niveles: un informe público puede incluir el objetivo de una instrucción y el método de control, mientras que el archivo restringido conserva el texto íntegro si contiene información que no debe difundirse. La apertura de la investigación no exige publicar indiscriminadamente cada conversación, y cualquier versión expurgada necesita indicar qué se ha omitido y cómo puede una persona autorizada examinar la evidencia relevante sin exponer a terceros.

Stahl et al. (2023), al abordar opciones para responder a desafíos éticos de la IA, relacionan la rendición de cuentas con auditorías y mecanismos que permiten examinar decisiones y atender sus consecuencias; trasladar ese criterio al expediente implica identificar quién puede pedir una revisión, qué información recibirá y qué facultad existe para corregir el resultado, en lugar de confiar en que una bitácora almacenada será consultada espontáneamente cuando aparezca un problema.

Si una referencia atribuida a un artículo no respalda el argumento, el registro debe permitir localizar el pasaje, retirar la atribución y revisar las partes que dependían de ella. Si un código cualitativo introduce un diagnóstico ausente, la corrección requiere volver al texto y ajustar la categoría, de modo que la trazabilidad funciona como infraestructura de reparación y no únicamente como un archivo destinado a demostrar que se siguieron pasos.

4.3.2. Reproducibilidad y evaluación ética del caso práctico

La auditoría del caso comienza por fijar el objeto revisado: el archivo base_analitica.csv del capítulo 3 contiene id, horas_semana y puntuacion, con sesenta filas y sesenta identificadores únicos; la comprobación ejecutada en este capítulo confirmó ausencia de celdas vacías, cumplimiento de los rangos definidos y coincidencia de la huella SHA-256, cuyo valor completo se conserva en el registro técnico, sin introducir datos nuevos ni sustituir las decisiones ya documentadas; se releyó el CSV y se recalculó Pearson mediante la fórmula explícita con aritmética decimal, obteniendo r = 0,8492297243905181, consistente con el valor publicado de 0,8492; esa segunda implementación constituye un control de cálculo sobre el mismo material y no una replicación con otra muestra, una revisión por un equipo independiente ni una prueba de que el procedimiento esté libre de cualquier error posible. La coincidencia numérica tiene un alcance preciso.

McDermott et al. (2021) examinan dificultades de reproducibilidad en investigación de aprendizaje automático para salud y señalan el acceso a datos y código como condiciones relevantes de revisión; aquí se retoma esa relación entre disponibilidad de materiales y posibilidad de comprobar procedimientos,

sin trasladar sus resultados al ámbito educativo ni afirmar que compartir un archivo garantiza la validez de sus supuestos o la legitimidad de su uso.

La revisión ética añade preguntas diferentes a las del cálculo: si las filas representan personas, si los textos proceden de entrevistas, si hubo autorizaciones declaradas sin respaldo y si las conclusiones exceden el diseño; en este caso, el expediente documenta generación procedimental y textos inventados, por lo que no se identificó tratamiento de información personal real dentro de esos insumos, aunque esta conclusión depende de la procedencia declarada y de los archivos concretos revisados.

Platzer y Reutterer (2021) proponen evaluar datos sintéticos atendiendo a fidelidad y privacidad mediante una estrategia con datos reservados, lo que ayuda a evitar la equivalencia automática entre síntesis y ausencia de riesgo; el ejercicio del libro tiene otra procedencia, pues fue generado desde reglas y distribuciones sin entrenar un sintetizador con registros personales, de modo que no corresponde atribuirle una evaluación de privacidad frente a una población original que nunca existió.

Para auditar el registro se utiliza una segunda instrucción, cuyo objeto es localizar inconsistencias entre lo propuesto, lo ejecutado y lo comunicado; la revisión se aplica a los estados documentados en los capítulos 3 y 4, sin solicitar un dictamen jurídico ni atribuir a un asistente la capacidad de conceder aprobaciones institucionales.

Prompt C4-P2 · Auditar el registro de decisiones

> Revisa el registro del caso simulado que adjunto. Identifica decisiones sin respaldo, cambios de versión no explicados, afirmaciones causales improcedentes y salidas de IA presentadas como verificadas sin evidencia de control. Clasifica los hallazgos y propone una corrección concreta para cada uno. No declares aprobaciones ni cumplimiento legal por tu cuenta.

Estado y salida. Instrucción reutilizable conservada como v1 el 27 de octubre de 2026; los hallazgos que siguen constituyen una revisión editorial del expediente disponible, acompañada por controles informáticos ejecutados sobre el CSV; no equivalen a evaluación por expertos externos ni dictamen legal.

Respuesta editorial aplicada al registro: la copia defectuosa conservó sesenta filas pero solo cincuenta y nueve códigos únicos después de eliminar una

ausencia, por lo que su coeficiente de 0,1511 no debe presentarse como resultado aceptado; los prompts C3-P1 y C3-P2 fueron propuestos y no intercambios externos ejecutados; los errores narrativos del capítulo anterior son ejemplos ficticios; la corrección aplicada mantiene esos estados visibles y reserva la etiqueta de ejecución para las operaciones respaldadas por código y salida; la misma lectura identifica límites que deben permanecer en el informe: no existe aprobación de comité ni validación por expertos externos, no se obtuvo consentimiento porque no hubo participantes y los fragmentos Q01–Q05 no pueden enlazarse a S001–S060; estas ausencias no se completan con fórmulas de cortesía editorial, y cualquier uso posterior de personas reales requeriría un protocolo propio antes de producir o incorporar esa evidencia.

Tabla 13.

Resultados de la auditoría C4-A01 y límites de cada control

Objeto	Control efectuado	Hallazgo	Alcance y límite
Archivo analítico	Lectura del CSV y SHA-256	60 filas, 60 id; huella idéntica al archivo del capítulo 3	Constata identidad del archivo revisado; no legitimidad de cualquier uso
Calidad estructural	Ausencias y rangos del diccionario	Sin vacíos; horas 0–40 y puntuación 0–100	Verifica reglas fijadas; no representatividad poblacional
Coeficiente	Fórmula explícita con Decimal sobre el CSV	r = 0,849229724390518…; coincide con 0,8492 publicado	Segunda implementación en el mismo entorno; sin muestra independiente
Procedencia	Revisión editorial de generación y materiales	Datos procedimentales y textos ficticios; sin vínculo entre Q y S	No constituye certificación jurídica ni prueba de equidad
Estados declarados	Contraste de texto, prompts, código y salida	Propuestas y errores ilustrativos identificados; sin aprobación inventada	Revisión editorial; no comité ni evaluación por expertos externos
Interpretación	Contraste con diseño y origen del caso	Asociación descriptiva conservada; causalidad rechazada	Sin eficacia educativa, generalización ni validación clínica

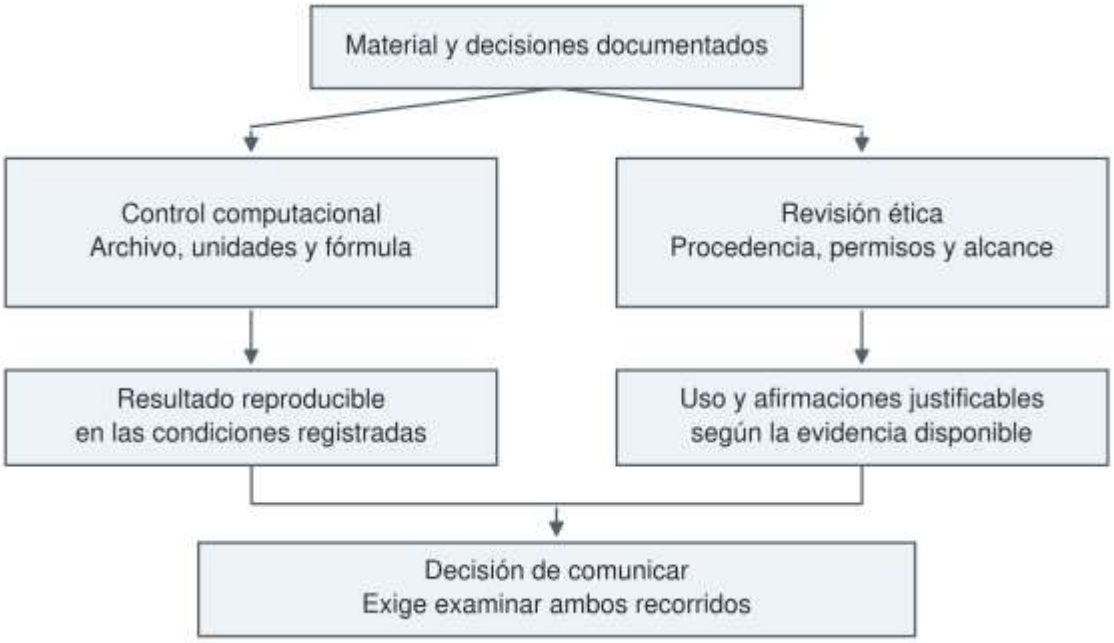
Las recomendaciones de Kapoor et al. (2024) para ciencia basada en aprendizaje automático ofrecen un marco consensuado de documentación y evaluación que refuerza la necesidad de explicitar elecciones y límites; su uso

aquí orienta la revisión del expediente y de las afirmaciones, sin convertir el seguimiento de una lista en una certificación de calidad, y sin afirmar que el caso sencillo de correlación constituye por sí mismo un estudio de entrenamiento y validación de modelos predictivos.

La reproducibilidad del coeficiente y la evaluación ética siguen recorridos relacionados, pero pueden llegar a conclusiones diferentes: una operación reproducible podría haber utilizado información sin autorización, mientras que un tratamiento autorizado podría producir un cálculo equivocado; la figura 6 hace visible esa separación para que el cierre de una revisión no se confunda con la aprobación automática de la otra.

Figura 6.

Controles distintos para el cálculo y la justificación del uso



Nota. Elaboración propia; los recorridos distinguen preguntas y evidencias; superar un control no garantiza el resultado del otro ni constituye una certificación general del estudio.

Ghassemi et al. (2021) advierten, en el contexto de la atención sanitaria, que las aproximaciones a la explicabilidad no garantizan por sí mismas seguridad o ausencia de sesgos y requieren validación rigurosa; la lección conceptual que se recupera para este capítulo consiste en no aceptar una explicación convincente

como sustituto del control pertinente, sin trasladar al caso educativo una estimación clínica ni atribuir equivalencia entre ambos ámbitos.

El dictamen del ejercicio puede expresarse de forma acotada: los controles ejecutados sostienen la consistencia del archivo y del coeficiente, mientras que la revisión editorial conserva su carácter sintético, rechaza inferencias causales y evita presentar autorizaciones inexistentes.

Quedan fuera de esa conclusión la representatividad poblacional, la equidad entre grupos, la utilidad de una intervención educativa y el cumplimiento jurídico de un proyecto real, aspectos para los que estos materiales no proporcionan evidencia suficiente. Si el caso se adapta con calificaciones o respuestas de estudiantes reales, debe abrirse una nueva revisión de procedencia, permisos y riesgos; conservar la plantilla no conserva las condiciones de la simulación, y el expediente tendrá que mostrar qué cambió, quién autorizó continuar y cómo se comprobó el resultado antes de incorporarlo al manuscrito científico.

Capítulo V: Comunicar hallazgos y producir conocimiento con rigor



Comunicar hallazgos y producir conocimiento con rigor

Comunicar una investigación exige mantener la relación entre la pregunta que la originó y las afirmaciones que pueden sostenerse al terminarla; cuando interviene la inteligencia artificial, esa relación necesita hacerse visible porque una herramienta puede mejorar la organización del manuscrito, proponer una explicación convincente o resumir una tabla sin que ninguna de esas operaciones añada evidencia al estudio, y el lector debe poder reconocer qué procede de los datos, qué interpretación asume el equipo y qué parte del proceso recibió asistencia.

El caso desarrollado a lo largo del libro llega a esta etapa con resultados calculados y decisiones documentadas: sesenta pares sintéticos sobre horas de estudio autónomo y puntuación, cinco fragmentos ficticios utilizados para enseñar codificación y un registro que conserva errores introducidos, correcciones y límites; cerrar el ejercicio consiste en comunicar ese recorrido con precisión, sin transformarlo en una investigación de campo ni convertir el coeficiente obtenido en una recomendación para estudiantes reales.

La escritura debe conservar el alcance que tenía la evidencia antes de ser reformulada.

5.1. Escritura científica y correspondencia entre afirmaciones y evidencia

Un manuscrito puede resultar claro en cada apartado y perder coherencia cuando se leen sus partes juntas: la introducción plantea una intervención, el método describe una encuesta, los resultados muestran una correlación y la conclusión anuncia una mejora; resolver ese desajuste requiere regresar a la pregunta y al diseño, porque añadir conectores o sustituir palabras repetidas no corrige la diferencia entre lo que se pretendía estudiar y lo que efectivamente se hizo; conviene comenzar la redacción de resultados desde un inventario breve de productos comprobables, con sus archivos y localizadores, antes de pedir una síntesis narrativa; en el caso del libro, ese inventario incluye la base analítica, los cálculos por dos vías y las incidencias documentadas, mientras que las

propuestas de prompt y las interpretaciones ilustrativas conservan un estado distinto que debe mantenerse cuando se preparen el resumen, las tablas y el cierre. Cada cifra necesita una función dentro del argumento.

Informar r = 0,8492 describe una asociación en el archivo simulado; acompañarlo de sesenta pares completos permite conocer la base del cálculo, y explicar que el generador incorporó una relación positiva impide atribuir el patrón a un descubrimiento empírico; en cambio, acumular decimales, añadir un valor de significación muy pequeño o llamar «robusto» al resultado sin precisar el control realizado puede aumentar la impresión de certeza sin mejorar lo que el lector sabe sobre su significado.

Jácome-Jácome et al. (2026), en el apartado dedicado a redacción científica, relacionan la calidad del texto con la correspondencia entre afirmaciones, método y evidencia, además de advertir que una reformulación puede alterar cautelas o conceptos; esta orientación resulta pertinente para revisar versiones de un manuscrito, donde una mejora de legibilidad debe conservar quién produjo el dato, qué se midió y con qué alcance se interpreta, aunque una frase más breve parezca estilísticamente preferible; la comprobación de referencias tiene dos objetos diferentes: identificar la obra y verificar que respalda la idea que se le atribuye; un título, un DOI y un año correctos permiten localizar una publicación, pero todavía queda por leer el pasaje pertinente y distinguir si presenta un resultado, una hipótesis o una recomendación de sus autores, de modo que la bibliografía no puede revisarse únicamente como un problema de formato al final de la escritura.

Walters y Wilder (2023) analizaron 636 referencias incluidas en 84 textos generados con GPT-3.5 y GPT-4 e identificaron tanto obras inventadas como errores sustantivos en referencias reales; su estudio muestra por qué la apariencia bibliográfica y la existencia del documento requieren controles separados, con resultados circunscritos a las versiones y condiciones examinadas, sin ofrecer una tasa de error aplicable a cualquier herramienta posterior; para una cita incorporada al capítulo, el recorrido de revisión conserva la identidad bibliográfica, el localizador del contenido y la lectura de la formulación ya insertada; ese tercer momento detecta cambios que pueden

pasar inadvertidos durante la paráfrasis, como convertir una percepción en un efecto, omitir que la muestra era no probabilística o trasladar a universitarios un resultado obtenido en otra población, aun cuando la referencia siga siendo auténtica y pertinente al tema general.

La tabla 14 aplica este criterio a seis enunciados del caso y separa la afirmación inicial de la evidencia y la decisión de redacción; las formulaciones incorrectas son ejemplos didácticos creados para esta revisión, no respuestas atribuidas a un servicio externo ni frases extraídas de una publicación real.

Tabla 14.

Mapa de afirmaciones y evidencia para comunicar el caso simulado

Afirmación inicial	Evidencia y localizador	Alcance	Revisión	Redacción final
A01. «Hay relación entre las variables»	C3, tabla 7; base_analitica.csv; r = 0,8492 y ρ = 0,8438	60 pares sintéticos	Precisar procedencia y tipo de relación	Se obtuvo asociación positiva en el archivo simulado
A02. «Una hora mejora 1,1985 puntos»	C3, tabla 7; pendiente del ajuste descriptivo	Diferencia del valor ajustado	Retirar causalidad	La pendiente ajustada fue de 1,1985 puntos por hora en esos datos
A03. «El modelo predice el rendimiento real»	Ajuste sobre los mismos 60 pares; sin evaluación externa	Sin evaluación predictiva externa	Rechazar el enunciado	No se evaluó la predicción en nuevos estudiantes
A04. «La conectividad explica la correlación»	C3, tabla 8; Q01–Q05 ficticios y sin enlace con S	Ejercicios separados	No cruzar como datos empíricos	Los fragmentos ilustran códigos, sin explicar el coeficiente
A05. «La IA mejoró el análisis»	Sin comparación con una ruta sin asistencia	No se midió eficacia comparativa	Retirar la ventaja atribuida	Se documentaron tareas asistidas y controles del procedimiento
A06. «El estudio recibió aprobación ética»	C4, tabla 13; no comité ni participantes reales	Simulación didáctica	Eliminar aprobación inexistente	Se realizó revisión editorial del caso sintético

El mapa permite revisar el manuscrito desde sus afirmaciones: una frase que no encuentra respaldo debe retirarse, acotarse o convertirse en una pregunta para otra investigación; completar la celda con una referencia sobre el mismo tema

no resuelve la ausencia de evidencia específica, del mismo modo que encontrar una explicación plausible para un error no autoriza a modificar silenciosamente la cifra publicada.

El resumen merece un control propio porque reúne, en poco espacio, decisiones que antes estaban distribuidas: diseño, origen de los datos, resultado central y alcance de la conclusión; al condensar un texto extenso, un asistente puede eliminar precisamente las palabras que fijan sus límites, de ahí que «simulado», «descriptivo» y «en este archivo» deban conservarse cuando su ausencia cambie la naturaleza del resultado, incluso si eso obliga a reducir otra parte menos informativa; en la comparación de Gao et al. (2023), los resúmenes generados a partir de títulos y nombres de revistas médicas presentaron una forma científica verosímil pese a contener datos producidos por el modelo; el hallazgo respalda la necesidad de revisar la procedencia de los resultados, porque reconocer la estructura de un resumen o encontrarlo convincente no equivale a comprobar que hubo participantes, mediciones y análisis que lo sustenten.

La variación de los párrafos también debe responder al argumento: una precisión breve puede fijar un límite antes de un desarrollo más extenso, mientras que un párrafo dedicado a comparar dos decisiones necesita espacio para explicar su relación. Conservar esa diversidad favorece una lectura natural, pero no añade participantes, no corrige una inferencia y no permite medir qué porcentaje del texto pertenece a una persona o a una herramienta. Los detectores tampoco sustituyen ese examen: Liang et al. (2023) encontraron clasificaciones erróneas de textos de autores no nativos de inglés en los sistemas que evaluaron, con implicaciones para la equidad de su uso.

El alcance lingüístico y experimental de ese estudio debe conservarse al citarlo, mientras que para este libro la consecuencia operativa es juzgar la integridad mediante fuentes, registros y correspondencia de afirmaciones, sin asignar al estilo una función de prueba sobre la autoría.

5.2. Autoría y declaración transparente del uso de inteligencia artificial

La autoría vincula una contribución con la capacidad de responder por el trabajo, aprobar lo que se publica y participar en la corrección de sus problemas; producir una parte del texto es una actividad relevante, pero su volumen no resuelve por sí solo la atribución de responsabilidad, y un acuerdo sobre firmas necesita distinguir quién formuló el problema, quién realizó o revisó el análisis y quién está en condiciones de defender las afirmaciones del manuscrito; la posición del Committee on Publication Ethics (2024) establece que las herramientas de IA no reúnen las condiciones para figurar como autoras y que su intervención debe declararse con transparencia; el criterio sitúa la responsabilidad del contenido en quienes firman el trabajo, incluyendo las partes asistidas, por lo que el reconocimiento de una herramienta no puede utilizarse para trasladarle la respuesta ante errores, conflictos de interés o incumplimientos editoriales.

Firmar también implica poder explicar una decisión concreta. Si una persona incorpora al artículo una tabla cuyo origen desconoce, su nombre no convierte esa tabla en un resultado defendible. Del mismo modo, distribuir tareas entre coautores exige acordar cómo se revisarán los componentes antes de aprobar el conjunto, porque conocer únicamente la sección que cada uno redactó puede dejar sin examinar contradicciones que aparecen entre el método, las cifras y la conclusión.

La taxonomía CRediT de la National Information Standards Organization (2022) ofrece catorce roles para describir contribuciones y señala expresamente que contributorship y autoría no son equivalentes; utilizarla ayuda a registrar actividades como conceptualización, curación de datos, análisis formal o revisión del texto, pero la asignación de un rol no determina automáticamente quién cumple los requisitos de autoría establecidos para una publicación concreta; en el expediente didáctico no corresponde completar una lista de investigadores con nombres supuestos ni asignar «validación» a especialistas que no participaron; sí puede conservarse la descripción de las operaciones realizadas y dejar que los autores responsables de la obra reconozcan sus contribuciones reales al

aprobar el manuscrito, con un acuerdo que identifique quién atenderá consultas y correcciones después de publicarlo.

Declarar «se utilizó IA» aporta demasiado poco si el lector necesita saber si la asistencia afectó únicamente a la sintaxis o intervino en el análisis; la descripción debe identificar la herramienta, las tareas y los controles disponibles, además de señalar una versión cuando conste en el registro, porque la omisión de ese dato y su invención producen problemas distintos y no deben resolverse de la misma manera.

En este ejercicio consta el uso de ChatGPT como asistente de redacción y organización, junto con preparación de instrucciones y código para análisis y gráficos; las operaciones numéricas publicadas se ejecutaron en Python, con NumPy 2.3.5, pandas 2.2.3 y SciPy 1.17.0 en el análisis del capítulo 3, y la revisión posterior conservó el archivo y contrastó Pearson mediante una fórmula decimal, sin que conste una evaluación independiente por expertos externos.

El identificador técnico exacto del modelo conversacional no está disponible en el expediente. Se registra esa ausencia para evitar que una etiqueta supuesta dé una apariencia de reproducibilidad que el historial no permite sostener. Cheng et al. (2025) proponen, en un comentario dirigido al ámbito de la simulación sanitaria, describir cómo se usaron herramientas generativas y cómo se trató su salida, además de preservar la contribución intelectual de los autores; su planteamiento ayuda a redactar una declaración por actividades, cuya ubicación y detalle deben ajustarse a la política del destino editorial, sin presentar aquellas recomendaciones como una demostración experimental de beneficios.

La tabla 15 organiza lo que puede afirmarse a partir de los registros conservados y distingue esa evidencia de las decisiones que requieren aprobación de quienes firmen la obra. Su utilidad consiste en impedir que una declaración bien redactada atribuya al equipo controles inexistentes o presente una muestra ilustrativa como una ejecución real.

Tabla 15.

Tareas y controles que puede recoger la declaración de uso de IA

Actividad	Registro disponible	Qué puede declararse	Límite del estado
Estructura y redacción	Texto y versiones editoriales de los capítulos	Asistencia de ChatGPT en organización y redacción	Sin identificador técnico exacto del modelo
Instrucciones seleccionadas	Prompts C3, C4 y C5 con estado indicado	Prompts propuestos y muestras o aplicaciones editoriales	No atribuir ejecución externa a toda instrucción
Cálculo y gráficos	Código, CSV, resultados y figuras originales	Ejecución en Python de las operaciones identificadas	No constituye evaluación empírica de la eficacia de IA
Revisión de fuentes	Registro bibliográfico y localizadores consultados	Consulta directa y controles documentales de citas	No equivale a revisión por especialistas independientes
Revisión y aprobación de autores	No consta aprobación final en el ejercicio	Describir únicamente actuaciones realizadas	Requiere revisión y aceptación de quienes firmen

Para preparar una declaración reutilizable se conserva el prompt C5-P2, cuyo último encargo depende de una política editorial concreta; como el ejercicio no se dirige aquí a una revista seleccionada, la muestra que sigue informa el uso documentado y no representa una declaración de conformidad con requisitos de envío todavía no identificados.

Prompt C5-P2 · Preparar una declaración transparente

Con el registro de herramientas y tareas que adjunto, redacta una declaración del uso de IA para el manuscrito. Identifica las tareas asistidas y las verificaciones humanas efectivamente documentadas. No atribuyas al equipo controles que no figuren en el registro ni afirmes que la IA es autora. Ajusta la declaración a la política editorial que aportaré.

Estado y entradas. Prompt propuesto, versión 1; muestra editorial del 27 de octubre de 2026 basada en los registros de los capítulos 3 y 4 y en la tabla 15, sin política de revista de destino aportada; la muestra no declara revisiones humanas ausentes ni cumplimiento de requisitos editoriales específicos.

Muestra editorial basada en el expediente: se utilizó ChatGPT para apoyar la estructura y redacción del material didáctico, proponer instrucciones y preparar código y gráficos; los resultados numéricos identificados como ejecutados

proceden de operaciones en Python sobre datos sintéticos y se contrastaron mediante los controles descritos, mientras que las fuentes se consultaron directamente y se revisó su correspondencia con las afirmaciones; no consta el identificador exacto del modelo ni una revisión independiente por especialistas, y la revisión disciplinar y aprobación final corresponden a quienes firmen el manuscrito.

Antes de utilizar este texto en un envío, los autores deberán comprobar que cada tarea los describe con fidelidad, añadir los datos de versión que realmente posean y aplicar la política vigente de la revista; también tendrán que revisar si se exige informar por separado sobre generación de imágenes, tratamiento de datos, traducción o análisis, porque una declaración general puede omitir intervenciones que el editor considere relevantes. La transparencia deja así una información que otro lector puede contrastar y que el equipo puede corregir cuando detecte una omisión.

5.3. Evaluación del aporte científico y de sus límites

El aporte de una investigación se aprecia por el problema que logra aclarar y por la evidencia que añade a la discusión de ese problema; producir más texto, explorar más combinaciones o completar un análisis con rapidez pueden cambiar la organización del trabajo, pero juzgar su contribución exige identificar qué se conoce mejor al terminar y qué alternativas siguen abiertas después de examinar los resultados; una contribución puede ser metodológica, como mostrar que una regla de depuración altera un resultado, o empírica, cuando aporta observaciones sobre un fenómeno; también puede consistir en construir una explicación o delimitar condiciones bajo las cuales una predicción funciona, y reconocer estas diferencias permite valorar un estudio sin exigirle descubrimientos que su diseño no estaba preparado para producir.

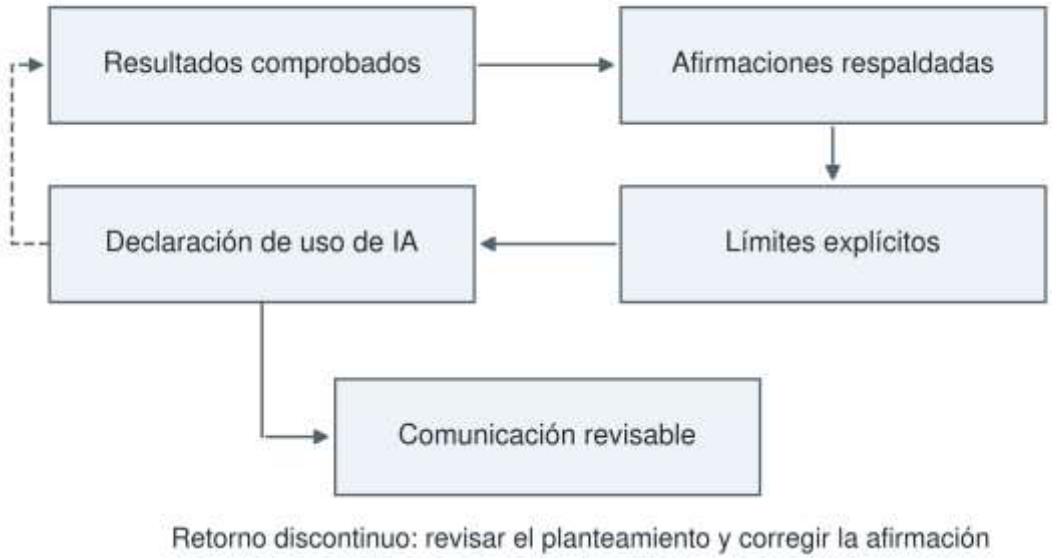
El ejercicio del libro aporta una demostración documentada del procedimiento. Krenn et al. (2022) examinan la comprensión científica asistida por IA y distinguen usos que amplían la exploración y la inspiración de la capacidad de comprender un fenómeno; esta discusión permite evaluar qué añade un sistema a la investigación más allá de emitir una respuesta acertada, porque una

predicción útil puede abrir preguntas sobre mecanismos que requieren un trabajo teórico y empírico adicional.

En el caso conductor se conoce la regla con la que se generaron las puntuaciones y se dispone de una ruta de análisis comprobada. Ese conocimiento permite enseñar cómo se conservan unidades, se detectan errores y se acotan inferencias, pero no transforma los sesenta registros en evidencia sobre estudiantes universitarios ni mide cuánto mejor habría trabajado un investigador sin asistencia, comparación que no se realizó. La figura 7 reúne las condiciones que conectan el resultado con una comunicación revisable y admite volver al planteamiento si la revisión encuentra una afirmación sin respaldo; la secuencia organiza el cierre editorial del caso y no representa una escala validada de calidad científica.

Figura 7.

De los resultados a una comunicación verificable



Nota. Elaboración propia; relación conceptual entre resultados, alcance y comunicación; la figura no representa datos empíricos ni una escala validada de calidad.

Recorrerla exige conservar el vínculo entre cada resultado y la decisión que autorizó su inclusión. Si se retira una frase causal, la conclusión debe revisarse también en el resumen y en cualquier material de divulgación; corregir solo el

cuerpo del capítulo dejaría circulando versiones incompatibles del mismo hallazgo.

5.3.1. Diferencias entre predicción, explicación y evidencia causal

Describir una asociación responde cómo varían conjuntamente dos medidas en el material disponible, mientras que predecir exige estimar un resultado todavía desconocido en condiciones definidas; una explicación intenta dar cuenta del patrón mediante relaciones conceptuales o mecanismos, y una afirmación causal añade una exigencia específica sobre qué cambiaría si se interviniera en una condición, por lo que compartir una fórmula o un conjunto de variables no vuelve intercambiables esos objetivos.

Laubach et al. (2021) distinguen descripción, predicción, asociación e inferencia causal y proponen definir la tarea antes de seleccionar el modelo; su argumento, desarrollado desde la biología y la epidemiología, sitúa el conocimiento previo y los supuestos sobre relaciones entre variables en la estrategia analítica, lo que ayuda a comprender por qué añadir un procedimiento estadístico no corrige una pregunta causal mal especificada. En el archivo sintético, la recta ajustada tiene intercepto 41,9000 y pendiente 1,1985 puntos por hora.

La pendiente describe una diferencia del valor ajustado a lo largo de las horas registradas, con $R^2 = 0,7212$ dentro de esos datos, pero la operación no incluyó una partición para evaluación predictiva ni observaciones externas, de modo que esos valores no acreditan la capacidad de anticipar resultados en una nueva cohorte. El tamaño de R^2 conserva el mismo límite que el archivo del que procede.

Una evaluación predictiva necesitaría fijar qué resultado se pretende anticipar, en qué momento estarán disponibles las variables y cómo se separarán desarrollo y evaluación; usar una calificación obtenida después del desenlace para predecirlo, por ejemplo, cambiaría el problema al introducir información que no existía cuando debía tomarse la decisión, aunque el ajuste resultante pareciera preciso dentro del conjunto analizado; si una futura investigación busca apoyar a estudiantes al inicio de un semestre, tendría que probar el

procedimiento con información disponible en ese momento y comparar su desempeño con una referencia pertinente; el criterio de utilidad incluiría además qué ocurre cuando el sistema se equivoca, porque detectar necesidades de apoyo y asignar una sanción requieren valorar consecuencias diferentes aun cuando compartan una medida de acierto.

Hofman et al. (2021) proponen integrar explicación y predicción en ciencias sociales computacionales mediante objetivos que pueden complementarse; su perspectiva permite evitar una elección artificial entre comprender y anticipar, siempre que el diseño haga explícito cómo se evalúa cada objetivo y no utilice un buen resultado en uno de ellos como evidencia suficiente del otro.

La explicación del coeficiente del caso tiene un punto de partida conocido: la puntuación se construyó con una relación lineal positiva respecto de las horas, un componente aleatorio y límites de escala; recuperar una asociación positiva permite revisar el procedimiento, mientras que atribuirla a motivación o a una estrategia de aprendizaje introduciría variables no observadas ni modeladas y cambiaría el objeto que se está explicando.

Del Giudice (2024) analiza la falacia de utilizar modelos optimizados para predecir como si ofrecieran por ese motivo una explicación adecuada y muestra que modelos con rendimiento semejante pueden sugerir relatos diferentes. Esta advertencia orienta la lectura de las salidas explicativas, donde la selección de variables, su importancia en un algoritmo o la sencillez de un relato deben examinarse frente al fenómeno que se pretende comprender.

Para una pregunta causal sobre horas de estudio y rendimiento sería necesario definir una intervención interpretable: aumentar el tiempo disponible, modificar una estrategia o proporcionar tutoría son cambios distintos que pueden tener efectos diferentes; además habría que considerar condiciones previas relacionadas con la dedicación y la puntuación, de modo que la frase «el efecto de estudiar más» necesita una definición operacional antes de convertirse en una estimación defendible; el libro de Huntington-Klein (2022) sitúa la identificación en la relación entre el proceso que genera los datos y la variación que permite responder una pregunta causal; aplicado al ejemplo, este planteamiento ayuda a distinguir el conocimiento de una regla simulada de la

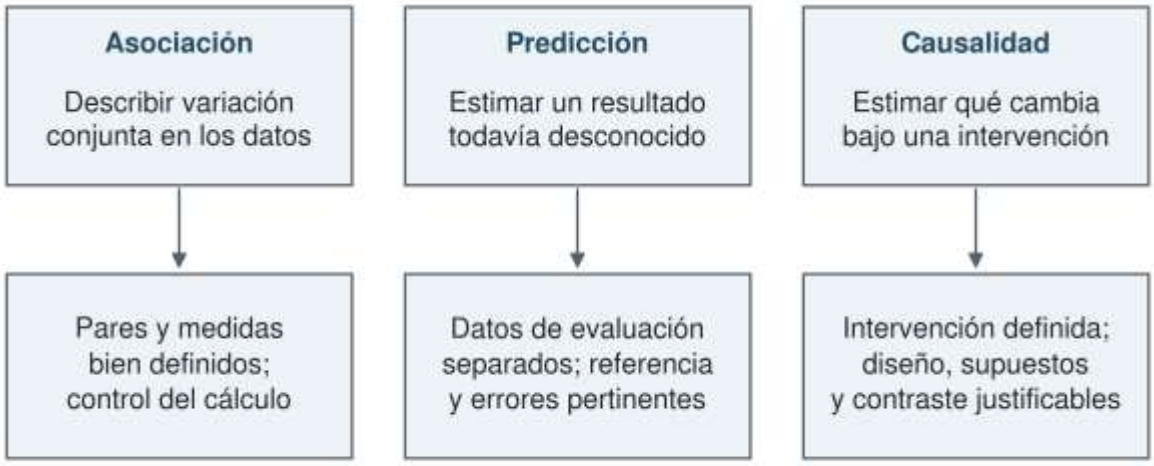
identificación de un efecto en personas, donde los mecanismos de selección, medición y asignación requieren evidencia y supuestos propios.

Conocer el generador ofrece una ventaja pedagógica, porque permite seguir la procedencia del patrón y revisar si el análisis recupera lo que se incorporó al construirlo. Esa ventaja desaparece cuando se interpreta la simulación como una observación del mundo, pues la validez de la regla para una población real no se obtiene al ejecutarla muchas veces ni al aumentar el número de filas inventadas.

La figura 8 compara las preguntas y la evidencia necesaria para tres tareas que podrían plantearse sobre el mismo tema; sus recorridos expresan condiciones de evaluación y no etapas obligatorias por las que todo estudio deba avanzar hasta alcanzar causalidad.

Figura 8.

Preguntas distintas y evidencia necesaria para responderlas



Nota. Elaboración propia; síntesis conceptual aplicada al razonamiento del caso; los recuadros indican requisitos de revisión, sin afirmar que todas esas evaluaciones se hayan realizado.

La comparación permite defender el valor de una descripción bien hecha: si el objetivo era examinar la consistencia de un procedimiento, no necesita presentarse como una intervención para resultar útil; el problema aparece

cuando la conclusión cambia de tarea sin aportar el diseño, los datos y la evaluación que exigiría ese cambio.

5.3.2. Ciencia abierta, colaboración y nuevas formas de descubrimiento

Abrir una investigación consiste en facilitar el examen y la reutilización de sus productos en condiciones comprensibles; publicar un CSV sin diccionario, una carpeta de código sin orden de ejecución o una figura sin procedencia puede aumentar la disponibilidad material y dejar intacta la dificultad de entender qué se hizo, por lo que la apertura requiere documentación, acceso adecuado y reglas explícitas para el uso de los materiales; la recomendación de la UNESCO (2021) vincula la ciencia abierta con transparencia, colaboración y acceso al conocimiento, al tiempo que admite restricciones justificadas por derechos, privacidad y confidencialidad; ese marco permite preparar materiales de distinto nivel de acceso y acompañar el código con especificaciones del entorno, sin equiparar apertura con difusión indiscriminada de registros personales o con una obligación idéntica para todos los contextos.

En el ejemplo sintético pueden conservarse la regla de generación, la semilla, el diccionario y el archivo analítico sin revelar experiencias de participantes que nunca existieron; para reutilizarlo, el lector necesitará conocer también el orden de las operaciones y qué errores fueron introducidos deliberadamente, porque una copia defectuosa aislada de ese contexto podría circular como si fuera un conjunto observado o un error accidental del análisis.

La disponibilidad empieza a ser útil cuando otra persona puede entender qué versión tiene delante. Un paquete complementario debería incluir una guía de lectura, los archivos de entrada y salida pertinentes y una indicación de las condiciones de reutilización. Las licencias del código y del texto no tienen que ser idénticas, y los materiales de terceros conservan sus propias condiciones, de modo que descargar una publicación de acceso abierto no autoriza a reasignar la propiedad de sus imágenes ni a distribuirla bajo cualquier licencia elegida para el proyecto.

Miedema (2022) examina la ciencia abierta como un cambio de organización e incentivos que afecta la relación entre investigación y sociedad; su planteamiento sitúa la colaboración y la revisión crítica junto a las infraestructuras, una relación relevante cuando se valora trabajo que no se resume en el número de artículos, como documentar una base, mantener código legible o permitir que otro equipo examine una decisión difícil; la colaboración necesita acuerdos concretos sobre contribuciones y versiones: si un grupo modifica una definición de variable y otro continúa con el diccionario anterior, reunir ambos análisis bajo el mismo título puede ocultar que respondieron preguntas distintas; conservar un registro común permite decidir si se corrige el cambio, se comparan versiones o se reconoce que los resultados pertenecen a objetos diferentes.

La revisión de Nosek et al. (2022) presenta la replicación como una práctica que confronta el conocimiento disponible y contribuye a su corrección, además de examinar los obstáculos culturales que dificultan realizarla; esa perspectiva da sentido a compartir materiales para que otros puedan cuestionarlos y no solo confirmarlos, sin reducir una discrepancia a un fallo personal de quien produjo o revisó el resultado; en este caso, volver a obtener Pearson con el mismo CSV es una comprobación computacional, y retirar un registro por vez examina una sensibilidad limitada del coeficiente; ninguno de esos controles constituye una réplica empírica con otra población, por lo que su documentación debe describir la operación y preservar la posibilidad de que una investigación posterior obtenga un patrón diferente bajo otras condiciones.

La IA puede ampliar el espacio de búsqueda científica cuando propone candidatos que después se someten a pruebas pertinentes. El valor del procedimiento depende tanto de la exploración como de la evaluación posterior, porque una lista extensa de hipótesis modifica lo que puede ensayarse, pero el número de propuestas no informa por sí solo cuántas resultarán correctas ni qué mecanismo explicará los resultados.

Jumper et al. (2021) evaluaron AlphaFold en CASP14 frente a estructuras de proteínas no divulgadas para los métodos participantes, lo que proporcionó un contraste externo a la generación de la predicción; el ejemplo muestra la importancia de especificar qué se predice y contra qué referencia se juzga,

conservando la diferencia entre estimar una estructura y establecer experimentalmente todos los procesos biológicos asociados con ella.

En descubrimiento de materiales, Merchant et al. (2023) combinaron filtrado mediante redes de grafos con evaluación de candidatos por cálculos de teoría del funcional de la densidad y distinguieron las estructuras calculadas de aquellas realizadas experimentalmente de manera independiente; el procedimiento ilustra cómo una búsqueda asistida necesita etapas de comprobación con alcance explícito, en lugar de llamar validación experimental a cada candidato identificado mediante simulación.

Estos ejemplos pertenecen a dominios con datos, infraestructuras y referencias de evaluación propios; su aportación transferible al caso educativo está en conservar la relación entre propuesta y prueba, mientras que trasladar sus métricas o resultados exigiría justificar otra correspondencia que el libro no ha establecido, especialmente si el objetivo posterior consiste en tomar decisiones sobre personas.

5.4. Cierre del caso y criterios transferibles a otras investigaciones

El cierre recupera una pregunta operativa: cómo describir y comprobar la relación entre horas de estudio autónomo y puntuación dentro de un conjunto sintético, preservando la procedencia de cada decisión; la unidad de análisis es la fila generada, identificada de S001 a S060, con horas entre 0 y 40 y puntuación entre 0 y 100, sin equivalencia con expedientes o estudiantes reclutados.

El procedimiento se inició con un generador de NumPy y semilla 20261026: las horas se obtuvieron de una distribución uniforme y la puntuación a partir de una relación lineal, error normal y recorte al rango, con redondeo a una decimal; conservar el orden de esas operaciones resulta necesario para reconstruir los valores, porque compartir únicamente la semilla dejaría sin especificar una parte sustantiva de la generación.

La copia de trabajo introdujo incidencias para enseñar su detección y la base depurada conservó sesenta códigos únicos y pares completos; las correcciones de unidades, ausencia y valor fuera de rango se justificaron desde el origen

sintético, y las comprobaciones publicadas obtuvieron Pearson r = 0,8492 y Spearman ρ = 0,8438, con coincidencia de Pearson mediante una segunda implementación sobre el CSV. Esos coeficientes describen el escenario construido.

La sensibilidad al retirar un registro por vez produjo valores de Pearson entre 0,8384 y 0,8649, intervalo que resume aquellas sesenta operaciones y no representa un intervalo de confianza.

Los fragmentos Q01–Q05, por su parte, permitieron mostrar cómo anclar un código en un texto ficticio y reconocer inferencias injustificadas, sin entrevistas reales ni una relación individual que permita combinarlos con las filas S.

El prompt C5-P1 se aplica como revisión editorial a un borrador defectuoso creado para el ejercicio: «El estudio demuestra que aumentar las horas de estudio mejora el rendimiento de los universitarios y que la IA permite analizar sus hábitos con mayor eficacia»; la frase reúne una atribución causal, una generalización poblacional y una comparación de eficacia que los materiales disponibles no sostienen.

Prompt C5-P1 · Revisar conclusiones frente a la evidencia

Te aporto una pregunta, una tabla de resultados calculados y el registro del caso simulado. Revisa cada afirmación del borrador de conclusiones y vincúlala con la evidencia disponible. Señala generalizaciones, causalidad no demostrada o cifras discordantes. Propón una versión breve que conserve la condición simulada del ejemplo. No añadas datos, citas ni resultados.

Estado y entradas. Prompt propuesto, versión 1; muestra editorial del 27 de octubre de 2026 basada en la pregunta del caso, los resultados de la tabla 7 y el registro C4-A01, con el borrador defectuoso reproducido en el párrafo anterior; no se presenta como una evaluación externa de un modelo identificado.

Respuesta editorial ilustrativa: retirar «demuestra» y la referencia a una mejora causal, conservar el rótulo de simulación y los coeficientes comprobados, y eliminar la ventaja atribuida a la IA porque no hubo comparación con un procedimiento sin asistencia; la decisión consiste en reemplazar la frase por una conclusión sobre el archivo y el control del procedimiento, sin añadir una investigación que no se realizó.

Conclusión delimitada del caso: en los sesenta pares sintéticos se obtuvo una asociación positiva entre horas de estudio autónomo y puntuación, con Pearson r = 0,8492 y Spearman ρ = 0,8438; el contraste de archivos, unidades y cálculo permitió documentar la consistencia del análisis, mientras que el origen simulado y la relación incorporada al generador limitan el resultado a una demostración metodológica, sin evidencia de efectos educativos en personas ni de una mejora atribuible al uso de IA.

Esa formulación resuelve la discrepancia entre el borrador y la evidencia, pero conserva las limitaciones que pertenecen al diseño; tampoco presenta el reconocimiento de errores como prueba de que se detectaron todos, porque cada control cubre un conjunto definido de problemas y la combinación de comprobaciones necesita seguir siendo examinable cuando alguien reutilice los materiales.

La tabla 16 identifica procedimientos reutilizables y las condiciones que habría que reconstruir en otra investigación; transferirlos requiere preservar su función de control y adaptar los criterios al nuevo objeto, en lugar de copiar rangos, identificadores o decisiones estadísticas que solo tenían sentido en este ejercicio.

Tabla 16.

Procedimientos transferibles y condiciones para su aplicación

Procedimiento	Utilidad	Condiciones necesarias	Límite	Evidencia que se conserva
Fijar pregunta y tarea	Alinear diseño y conclusión	Definir unidad, población, objetivo y resultado	La pregunta del caso no se copia automáticamente	Protocolo y cambios justificados
Revisar el diccionario	Evitar errores de escala o unidad	Definiciones pertinentes al nuevo estudio	Los rangos 0–40 y 0–100 pertenecen al ejemplo	Diccionario, unidades y procedencia
Trazar transformaciones	Reconstruir el análisis	Origen autorizado y reglas explícitas	No recuperar un dato real desde una fuente inventada	Entrada, salida y motivo del cambio
Contrastar cálculos	Detectar diferencias de cálculo	Procedimiento adecuado y	Coincidir no demuestra	Código, entorno y

		control con alcance definido	validez de los supuestos	salida de control
Revisar afirmaciones	Ajustar lenguaje a evidencia	Localizador específico y examen del diseño	No convertir una asociación en efecto por reformular	Mapa afirmación–evidencia y versión revisada
Preparar materiales	Facilitar examen y reutilización	Permisos, documentación y acceso apropiado	No publicar datos restringidos sin garantías	Inventario, condiciones de acceso y versiones

Nota. Elaboración propia; orientaciones metodológicas derivadas del recorrido documentado, sin validación como instrumento ni equivalencia entre disciplinas, poblaciones o diseños.

En una investigación con datos personales, el primer cambio sería la procedencia y autorización de los insumos, con las garantías correspondientes antes de utilizar herramientas externas; en un estudio cualitativo, la revisión tendría que conservar contexto y decisiones interpretativas, mientras que una aplicación predictiva necesitaría un diseño de evaluación propio, de modo que la posibilidad de reutilizar una bitácora no convierte esos proyectos en variantes equivalentes del mismo análisis; para trasladar el procedimiento, resulta útil empezar por una decisión que el nuevo estudio necesite defender y determinar qué evidencia permitiría revisarla: una exclusión puede requerir el criterio de elegibilidad y su aplicación al registro, una transformación exige unidades y código, y una conclusión necesita datos o fuentes que respalden exactamente su formulación; esta orientación concentra la documentación en operaciones con consecuencias sobre el resultado.

La tabla 17 reúne los materiales que permiten cerrar la comunicación del ejercicio y distingue su estado actual de actuaciones que dependerían de una publicación posterior; el inventario conserva lo producido sin atribuir al caso un depósito público, una licencia editorial definitiva o una aprobación humana que no constan en el expediente.

Tabla 17.

Inventario de cierre y estado de los materiales del ejercicio

Material	Contenido conservado	Estado documentado	Uso posterior
Base y diccionario	60 filas S001–S060; horas_semana y puntuacion	Archivo sintético conservado, sin modificación en el cierre	Reutilizar con su procedencia explícita
Generación y análisis	Semilla, código, entorno y resultados del C3	Operaciones identificadas ejecutadas y registradas	Reconstruir el orden de ejecución
Auditoría ética	C4-A01; controles y límites declarados	Revisión técnica y editorial; sin dictamen legal o comité	Revisar de nuevo si cambian los insumos
Comunicación del caso	Mapa de afirmaciones, conclusión y declaración base	Elaboración editorial del C5	Revisión disciplinar y aprobación de autores
Acceso y publicación	Inventario de materiales y condiciones necesarias	No consta depósito público ni licencia editorial definitiva	Elegir destino, permisos y condiciones antes de difundir

El cierre del libro deja un criterio para cada afirmación: conservar su procedencia, el control recibido y sus límites de uso; cuando una versión posterior modifique cifras o interpretaciones, el registro deberá explicar el cambio y la corrección tendrá que llegar también al resumen y a los materiales derivados para que otras personas puedan examinarla.

Referencias Bibliográficas

Aeon, B., Faber, A., & Panaccio, A. (2021). Does time management work? A meta-analysis. *PLOS ONE, 16*(1), Artículo e0245066. https://doi.org/10.1371/journal.pone.0245066

Afzaal, M., & Nouri, J. (2024). A systematic review of software for learning analytics in higher education. *International Journal of Emerging Technologies in Learning, 19*(7), 17–43. https://doi.org/10.3991/ijet.v19i07.50313

Aucapiña-Tubón, A. P., Rodríguez-Loor, J. E., Bombón-Chico, S. H., & Salazar-Cedeño, J. J. (2026). Uso de la IA en la educación: Desafíos éticos y consecuencias sociales en Ecuador. *Innova Science Journal, 4*(2), 569–579. https://doi.org/10.63618/omd/isj/v4/n2/296

Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile* (NIST AI 600-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.600-1

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? En *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns, 2*(2), 100205. https://doi.org/10.1016/j.patter.2021.100205

Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1), Artículo 188. https://doi.org/10.1145/3449287

Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity, 56*(3), 1391–1412. https://doi.org/10.1007/s11135-021-01182-y

Cheng, A., Calhoun, A., & Reedy, G. (2025). Artificial intelligence-assisted academic writing: Recommendations for ethical use. *Advances in Simulation, 10*, Artículo 22. https://doi.org/10.1186/s41077-025-00350-6

Committee on Publication Ethics. (2024). *Authorship and AI tools* (Versión 1). https://doi.org/10.24318/cCVRZBms

Congreso Constituyente Democrático. (1993). *Constitución Política del Perú*. Artículo 2, numeral 6. Edición del Congreso de la República, noviembre

de 2022, p. 12. https://www3.congreso.gob.pe/Docs/biblioteca/Constitucion/CP-ED-Nov2022/files/basic-html/page12.html

Congreso de la República del Perú. (2011). *Ley N.º 29733, Ley de Protección de Datos Personales*. Diario Oficial El Peruano, 3 de julio de 2011. Texto actualizado disponible en la Superintendencia del Mercado de Valores. https://www.smv.gob.pe/Uploads/Ley_29733_vigente_2025.pdf

Congreso de la República del Perú. (2023). *Ley N.º 31814, Ley que promueve el uso de la inteligencia artificial en favor del desarrollo económico y social del país*. Diario Oficial El Peruano, 5 de julio de 2023. https://busquedas.elperuano.pe/dispositivo/NL/2192926-1

Del Giudice, M. (2024). The prediction-explanation fallacy: A pervasive problem in scientific applications of machine learning. *Methodology, 20*(1), 22–46. https://doi.org/10.5964/meth.11235

De Paoli, S. (2024). Performing an inductive thematic analysis of semi-structured interviews with a large language model: An exploration and provocation on the limits of the approach. *Social Science Computer Review, 42*(4), 997–1019. https://doi.org/10.1177/08944393231220483

Floridi, L. (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press. https://doi.org/10.1093/oso/9780198883098.001.0001

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *npj Digital Medicine, 6*, Artículo 75. https://doi.org/10.1038/s41746-023-00819-6

Gartlehner, G., Kahwati, L., Hilscher, R., Thomas, I., Kugley, S., Crotty, K., Viswanathan, M., Nussbaumer-Streit, B., Booth, G., Erskine, N., Konet, A., & Chew, R. (2024). Data extraction for evidence synthesis using a large language model: A proof-of-concept study. *Research Synthesis Methods, 15*(4), 576–589. https://doi.org/10.1002/jrsm.1710

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé, H., III, & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM, 64*(12), 86–92. https://doi.org/10.1145/3458723

Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health, 3*(11), e745–e750. https://doi.org/10.1016/S2589-7500(21)00208-9

Gusenbauer, M. (2022). Search where you will find most: Comparing the disciplinary coverage of 56 bibliographic databases. *Scientometrics, 127*(5), 2683–2745. https://doi.org/10.1007/s11192-022-04289-7

Haddaway, N. R., Grainger, M. J., & Gray, C. T. (2022). Citationchaser: A tool for transparent and efficient forward and backward citation chasing in systematic searching. *Research Synthesis Methods, 13*(4), 533–545. https://doi.org/10.1002/jrsm.1563

Haddaway, N. R., Rethlefsen, M. L., Davies, M., Glanville, J., McGowan, B., Nyhan, K., & Young, S. (2022). A suggested data structure for transparent and repeatable reporting of bibliographic searching. *Campbell Systematic Reviews, 18*(4), Artículo e1288. https://doi.org/10.1002/cl2.1288

Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., & Yarkoni, T. (2021). Integrating explanation and prediction in computational social science. *Nature, 595*(7866), 181–188. https://doi.org/10.1038/s41586-021-03659-0

Hooker, S. (2021). Moving beyond "algorithmic bias is a data problem". *Patterns, 2*(4), 100241. https://doi.org/10.1016/j.patter.2021.100241

Hosseini, M., Rasmussen, L. M., & Resnik, D. B. (2024). Using AI to write scholarly publications. *Accountability in Research, 31*(7), 715–723. https://doi.org/10.1080/08989621.2023.2168535

Hosseini, M., Resnik, D. B., & Holmes, K. (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics, 19*(4), 449–465. https://doi.org/10.1177/17470161231180449

Huntington-Klein, N. (2022). *The effect: An introduction to research design and causality*. Chapman and Hall/CRC. https://doi.org/10.1201/9781003226055

Jácome-Jácome, J. P. (2026). La inteligencia artificial como agente de innovación pedagógica orientada a la personalización del aprendizaje en educación superior. *Innova Science Journal, 4*(1), 1–19. https://doi.org/10.63618/omd/isj/v4/n1/205

Jácome-Jácome, J. P., Alarcón-Anco, R. J., Chávez-Rodríguez, M. A., Espinel-Bermeo, A. C., Ávila-Guaman, D. E., Sornoza-Bravo, J. C., Pico-Chavarría, V. F., Sánchez-Vásquez, M. E., Zambrano-Zambrano, A. B., & Sánchez-Vásquez, M. C. (2026). *Inteligencia artificial aplicada a la educación, la investigación y las organizaciones: Casos, herramientas y perspectivas para América Latina.* Oriente-Manabí Editorial. https://doi.org/10.63618/omeditorial/l33

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). *An introduction to statistical learning: With applications in Python*. Springer. https://doi.org/10.1007/978-3-031-38747-0

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Juna-Pozo, L. P. (2025). Percepciones sobre la inteligencia artificial analizadas con componentes principales y regresión multinomial. *Innova Science Journal, 3*(E1), 289–302. https://doi.org/10.63618/omd/isj/v3/nE1/194

Kapoor, S., Cantrell, E. M., Peng, K., Pham, T. H., Bail, C. A., Gundersen, O. E., Hofman, J. M., Hullman, J., Lones, M. A., Malik, M. M., Nanayakkara, P., Poldrack, R. A., Raji, I. D., Roberts, M., Salganik, M. J., Serra-Garcia, M., Stewart, B. M., Vandewiele, G., & Narayanan, A. (2024). REFORMS: Consensus-based recommendations for machine-learning-based science. *Science Advances, 10*(18), eadk3452. https://doi.org/10.1126/sciadv.adk3452

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns, 4*(9), Artículo 100804. https://doi.org/10.1016/j.patter.2023.100804

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Artículo 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khraisha, Q., Put, S., Kappenberg, J., Warraitch, A., & Hadfield, K. (2024). Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Research Synthesis Methods, 15*(4), 616–626. https://doi.org/10.1002/jrsm.1715

Krenn, M., Pollice, R., Guo, S. Y., Aldeghi, M., Cervera-Lierta, A., Friederich, P., dos Passos Gomes, G., Häse, F., Jinich, A., Nigam, A., Yao, Z., & Aspuru-Guzik, A. (2022). On scientific understanding with artificial intelligence. *Nature Reviews Physics, 4*(12), 761–769. https://doi.org/10.1038/s42254-022-00518-3

Lakens, D. (2022). Sample size justification. *Collabra: Psychology, 8*(1), Artículo 33267. https://doi.org/10.1525/collabra.33267

Laubach, Z. M., Murray, E. J., Hoke, K. L., Safran, R. J., & Perng, W. (2021). A biologist's guide to model selection and causal inference. *Proceedings of the Royal Society B: Biological Sciences, 288*(1943), Artículo 20202815. https://doi.org/10.1098/rspb.2020.2815

Lee, K. J., Tilling, K. M., Cornish, R. P., Little, R. J. A., Bell, M. L., Goetghebeur, E., Hogan, J. W., & Carpenter, J. R. (2021). Framework for the treatment and reporting of missing data in observational studies: The Treatment And Reporting of Missing data in Observational Studies framework. *Journal of Clinical Epidemiology, 134*, 79–88. https://doi.org/10.1016/j.jclinepi.2021.01.008

Lee, V. V., van der Lubbe, S. C. C., Goh, L. H., & Valderas, J. M. (2024). Harnessing ChatGPT for thematic analysis: Are we ready? *Journal of Medical Internet Research, 26*, Artículo e54974. https://doi.org/10.2196/54974

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns, 4*(7), Artículo 100779. https://doi.org/10.1016/j.patter.2023.100779

Lones, M. A. (2024). Avoiding common machine learning pitfalls. *Patterns, 5*(10), Artículo 101046. https://doi.org/10.1016/j.patter.2024.101046

Mahmoudi, H., Chang, D., Lee, H., Ghaffarzadegan, N., & Jalali, M. S. (2025). Critical assessment of large language models' (ChatGPT) performance in data extraction for systematic reviews: Exploratory study. *JMIR AI, 4*, Artículo e68097. https://doi.org/10.2196/68097

Mancheno-Oviedo, M. P., Garcés-Yanzapanta, S. D., Pinto-Almeida, G. D., & Rodríguez-Loor, J. E. (2026). Uso de la inteligencia artificial en el desarrollo del pensamiento crítico en estudiantes de bachillerato. *Innova Science Journal, 4*(3), 215–228. https://doi.org/10.63618/omd/isj/v4/n3/341

Mathis, W. S., Zhao, S., Pratt, N., Weleff, J., & De Paoli, S. (2024). Inductive thematic analysis of healthcare qualitative interviews using open-source large language models: How does it compare to traditional methods? *Computer Methods and Programs in Biomedicine, 255*, Artículo 108356. https://doi.org/10.1016/j.cmpb.2024.108356

McDermott, M. B. A., Wang, S., Marinsek, N., Ranganath, R., Foschini, L., & Ghassemi, M. (2021). Reproducibility in machine learning for health research: Still a ways to go. *Science Translational Medicine, 13*(586), eabb1655. https://doi.org/10.1126/scitranslmed.abb1655

McKinney, W. (2022). *Python for data analysis: Data wrangling with pandas, NumPy, and Jupyter* (3.ª ed.). O'Reilly Media. https://wesmckinney.com/book/

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), Artículo 115. https://doi.org/10.1145/3457607

Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. *Nature, 624*(7990), 80–85. https://doi.org/10.1038/s41586-023-06735-9

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature, 627*(8002), 49–58. https://doi.org/10.1038/s41586-024-07146-0

Miedema, F. (2022). *Open science: The very idea*. Springer. https://doi.org/10.1007/978-94-024-2115-6

Morgan, D. L. (2023). Exploring the use of artificial intelligence for qualitative data analysis: The case of ChatGPT. *International Journal of Qualitative Methods, 22*, 1–10. https://doi.org/10.1177/16094069231211248

National Information Standards Organization. (2022). *CRediT, contributor roles taxonomy (ANSI/NISO Z39.104-2022)*. https://doi.org/10.3789/ansi.niso.z39.104-2022

National Library of Medicine. (2026, 30 de septiembre). *PubMed user guide.* https://pubmed.ncbi.nlm.nih.gov/help/

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology, 73*, 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

OECD. (2023). *Artificial intelligence in science: Challenges, opportunities and the future of research.* OECD Publishing. https://doi.org/10.1787/a8d820bd-en

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA

2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine, 18*(3), Artículo e1003583. https://doi.org/10.1371/journal.pmed.1003583

Peralta Palazón, T. (2026). Conocimiento o apariencia: El estatuto epistémico de la inteligencia artificial generativa. En A. Escudero-Nahón (Ed.), *Inteligencia artificial: Experiencias y reflexiones sobre la investigación científica* (pp. 312–324). Editorial Transdigital. https://doi.org/10.56162/transdigitalbc12.18

Peters, M. D., Godfrey, C., McInerney, P., Khalil, H., Larsen, P., Marnie, C., Pollock, D., Tricco, A. C., & Munn, Z. (2022). Best practice guidance and reporting items for the development of scoping review protocols. *JBI Evidence Synthesis, 20*(4), 953–968. https://doi.org/10.11124/JBIES-21-00242

Pineau, J., Vincent-Lamarre, P., Sinha, K., Larivière, V., Beygelzimer, A., d'Alché-Buc, F., Fox, E., & Larochelle, H. (2021). Improving reproducibility in machine learning research (A report from the NeurIPS 2019 reproducibility program). *Journal of Machine Learning Research, 22*(164), 1–20. https://jmlr.org/papers/v22/20-303.html

Platzer, M., & Reutterer, T. (2021). Holdout-based empirical assessment of mixed-type synthetic data. *Frontiers in Big Data, 4*, 679939. https://doi.org/10.3389/fdata.2021.679939

Presidencia de la República del Perú. (2024). *Decreto Supremo N.º 016-2024-JUS que aprueba el Reglamento de la Ley N.º 29733, Ley de Protección de Datos Personales*. Diario Oficial El Peruano, Separata Especial, 30 de noviembre de 2024. https://www.sat.gob.pe/TransparenciaV3/Portals/0/Docs/NormasOtrasEntidades/DecretoSupremo016-2024-JUS_ReglamentoLey29733_LeyProteccionDatosPersonales.pdf

Presidencia del Consejo de Ministros. (2025). *Decreto Supremo N.º 115-2025-PCM que aprueba el Reglamento de la Ley N.º 31814*. Diario Oficial El Peruano, 9 de septiembre de 2025. https://busquedas.elperuano.pe/dispositivo/NL/2436426-1

Resnik, D. B., & Hosseini, M. (2025). The ethics of using artificial intelligence in scientific research: New guidance needed for a new tool. *AI and Ethics, 5*(2), 1499–1521. https://doi.org/10.1007/s43681-024-00493-8

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews, 10*, Artículo 39. https://doi.org/10.1186/s13643-020-01542-z

Rittman, M. (2025, 29 de enero). *Retraction Watch retractions now in the Crossref API.* Crossref. https://www.crossref.org/blog/retraction-watch-retractions-now-in-the-crossref-api/

Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical Care, 27*, Artículo 75. https://doi.org/10.1186/s13054-023-04380-2

Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., & Bengio, Y. (2021). Toward causal representation learning. *Proceedings of the IEEE, 109*(5), 612–634. https://doi.org/10.1109/JPROC.2021.3058954

scikit-learn. (2026). Common pitfalls and recommended practices. En *scikit-learn 1.9.1 documentation*. Recuperado el 26 de octubre de 2026, de https://scikit-learn.org/stable/common_pitfalls.html

SciPy. (2026). pearsonr. En *SciPy v1.18.0 manual*. Recuperado el 26 de octubre de 2026, de https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html

Stahl, B. C. (2021). *Artificial intelligence for a better future: An ecosystem perspective on the ethics of AI and emerging digital technologies*. Springer. https://doi.org/10.1007/978-3-030-69978-9

Stahl, B. C., Schroeder, D., & Rodrigues, R. (2023). *Ethics of artificial intelligence: Case studies and options for addressing ethical challenges*. Springer. https://doi.org/10.1007/978-3-031-17040-9

Torres-Vela, S. A., Barbachan-Ruales, E. A., Aponte-Limaco, J. J., Mendoza-Avendaño, E. A., & Yana-Quispilaya, D. D. (2025). *Formación investigativa, tecnodigitalización docente e innovación metodológica.* Oriente-Manabí Editorial. https://doi.org/10.63618/omeditorial/l4

UNESCO. (2021). *Recomendación de la UNESCO sobre la ciencia abierta*. https://www.unesco.org/en/legal-affairs/recommendation-open-science

UNESCO. (2026, 16 de enero). *Guidance for generative AI in education and research.* https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour, 8*(12), 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

Valencia-Galarza, B. N., Alarcón-Anco, R. J., Cedeño-Zambrano, L. P., Valdivieso-Mora, M. C., Manzaba-Gilces, I. del J., Mieles-Valdivieso, C. A., Bucheli-Valdivieso, C. S., Segovia-García, C. C., Bravo-Zambrano, Y.

D., & Cevallos-Vera, N. G. (2026). *Educación superior inteligente: Transformación digital, inteligencia artificial y aprendizaje adaptativo en la formación universitaria.* Oriente-Manabí Editorial. https://doi.org/10.63618/omeditorial/l24

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence, 3*(2), 125–133. https://doi.org/10.1038/s42256-020-00287-7

van Dis, E. A. M., Bollen, J., van Rooij, R., Zuidema, W., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature, 614*(7947), 224–226. https://doi.org/10.1038/d41586-023-00288-7

Varoquaux, G., & Cheplygina, V. (2022). Machine learning for medical imaging: Methodological failures and recommendations for the future. *npj Digital Medicine, 5*, Artículo 48. https://doi.org/10.1038/s41746-022-00592-y

Velasco-Holguin, N. P., Alarcón-Anco, R. J., Alcívar-Cuenca, P. E., García-Lino, M. E., Ramos-Zapata, A. L., Araujo-Criollo, M. D., Dávila-Alva, E., Chang-Gómez, M. A., Reategui-Noriega, E. E., & Guerrero-Zambrano, M. F. (2026). *Educación basada en datos: Learning Analytics para mejorar el rendimiento académico y la gestión universitaria.* Oriente-Manabí Editorial. https://doi.org/10.63618/omeditorial/l25

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. *Scientific Reports, 13*, Artículo 14045. https://doi.org/10.1038/s41598-023-41032-5

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., Anandkumar, A., Bergen, K., Gomes, C. P., Ho, S., Kohli, P., Lasenby, J., Leskovec, J., Liu, T. Y., Manrai, A., … Zitnik, M. (2023). Scientific discovery in the age of artificial intelligence. *Nature, 620*(7972), 47–60. https://doi.org/10.1038/s41586-023-06221-2

Waskom, M. L. (2021). seaborn: Statistical data visualization. *Journal of Open Source Software, 6*(60), Artículo 3021. https://doi.org/10.21105/joss.03021

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19*, Artículo 26. https://doi.org/10.1007/s40979-023-00146-z

Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). *R for data science: Import, tidy, transform, visualize, and model data* (2.ª ed.). O'Reilly Media. https://r4ds.hadley.nz/

Resumen

Este libro examina la incorporación de la inteligencia artificial en la investigación científica mediante un recorrido que articula decisiones metodológicas, herramientas y responsabilidad académica; su propósito es orientar un uso crítico que preserve la correspondencia entre preguntas, evidencia y conclusiones, atendiendo a condiciones de acceso y formación relevantes para América Latina; la obra desarrolla la búsqueda y verificación de fuentes, la preparación y análisis de datos, la protección de información y la comunicación de hallazgos, con instrucciones seleccionadas, registros de decisiones y controles cuyo alcance se explica; un caso didáctico con sesenta registros sintéticos permite mostrar la depuración de datos, el contraste de cálculos y la revisión de afirmaciones, acompañado por fragmentos cualitativos ficticios que conservan su carácter ilustrativo; el análisis distingue asociación, predicción y causalidad, y examina la autoría, la declaración del uso de herramientas y las condiciones de una ciencia abierta compatible con derechos y permisos; el aporte consiste en ofrecer procedimientos revisables y criterios transferibles, sin presentar la simulación como evidencia sobre estudiantes reales ni atribuir eficacia comparativa a la asistencia tecnológica; se propone una práctica investigativa que conserve la procedencia de los materiales y permita comprender, cuestionar y corregir los resultados comunicados a la comunidad científica.

Palabras clave: inteligencia artificial; investigación científica; integridad científica; reproducibilidad; escritura académica.

Abstract

This book examines the integration of artificial intelligence into scientific research through an approach connecting methodological decisions, tools, and academic responsibility; its purpose is to guide critical use that preserves correspondence among questions, evidence, and conclusions while addressing access and training conditions relevant to Latin America; the book covers source searching and verification, data preparation and analysis, information protection, and the communication of findings through selected instructions, decision records, and checks with explicit scopes; a teaching case involving sixty synthetic records demonstrates data cleaning, computational crosschecking, and the review of claims, accompanied by fictional qualitative excerpts whose illustrative status remains explicit; the analysis distinguishes association, prediction, and causality while examining authorship, disclosure of tool use, and conditions for open science consistent with rights and permissions; its contribution lies in providing reviewable procedures and transferable criteria without presenting simulation as evidence about actual students or attributing comparative effectiveness to technological assistance; the proposed research practice preserves the provenance of materials and enables readers to understand, question, and correct reported results, with methodological boundaries and uncertainties clearly stated, emphasizing that the usefulness of artificial intelligence depends on methodological judgment and documented verification rather than textual fluency or the volume of generated content.

Keywords: artificial intelligence; scientific research; research integrity; reproducibility; academic writing.

ISBN: 978-9907-9620-3-1

9 789907 962031

